ANALYSIS OF AN M/M/1 QUEUE WITH CUSTOMER INTERJECTION

by

Alireza Aliakbar Chavoushi

Submitted in partial fulfilment of the requirements
for the degree of Master of Applied Science

at

Dalhousie University
Halifax, Nova Scotia
June, 2010

DALHOUSIE UNIVERSITY

Faculty of Engineering

Industrial Engineering

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled “ANALYSIS OF AN M/M/1 QUEUE WITH CUSTOMER INTERJECTION” by Alireza Aliakbar Chavoushi in partial fulfillment of the requirements for the degree of Master of Applied Science.

Dated: ____________________

Supervisor: ____________________________

Dr. Qi-Ming He

Examiners:

____

Dr. Pemberton Cyrus

____

Dr. Carl-Louis Sandblom

____

Dr. Jacek Ilow

DALHOUSIE UNIVERSITY

DATE: June 24, 2010

AUTHOR: Alireza Aliakbar Chavoushi

TITLE: Analysis of an M/M/1 Queue with Customer Interjection

DEPARTMENT OR SCHOOL: Industrial Engineering

DEGREE: Master of Applied Science CONVOCATION: October YEAR: 2010

____

Signature of Author

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to express my gratitude to my supervisor, Dr. Qi-Ming He, who, during the years in which this endeavour lasted, provided me continuous guidance and encouragement. Next, I would like to thank my committee members, Dr. Pemberton Cyrus, Dr. Carl-Louis Sandblom and Dr. Jacek Ilow for taking time and efforts to review my thesis. I am also thankful to my professors and friends at Dalhousie University Industrial Engineering Department for their contributions and help on this thesis.

I am grateful to my parents, Mahboubeh Fard Hoseini and Fazel Aliakbar Chavoushi, and my sisters, Raheleh and Reyhaneh, for believing in me, and for their continuous support and encouragement during the creation of this work.

Abstract

In our daily life, we often experience waiting in a queue to receive some kind of service. Some customers do not join the queue at the end like other normal customers, and try to cut in the queue hoping to have a shorter waiting time and a higher level of satisfaction. This behaviour is called *customer interjection*. Some of these customers only try to cut in queue, while some others try to find excuses for interjection. For instance, the first-come-first-served (FCFS) service discipline is usually assumed in public places like restaurants, banks, airports, and supermarkets. However, customer interjections can still be seen in these places. In telecommunications networks, to test the efficiency of transmission, artificial packages are inserted into the normal traffic in a random manner. These interjections can affect the waiting time of other customers in queue. Such interjections may reduce the waiting time of interjecting customers, but increase the waiting time and dissatisfaction of others.

In this work, an M/M/1 queueing system with customer interjection is investigated. The arrival of customers to the system is assumed to be a Poisson process with arrival rate λ. The service times for customers are independent and identically distributed random variables with an exponential distribution with rate μ. Customers are dispersed into normal customers and interjecting customers. A normal customer joins the queue at the end, and an interjecting customer tries to cut in the queue and occupy a position as close to the head of the queue as possible. Two parameters are introduced to describe the interjection behaviour: the percentage of customers interjecting and the tolerance level of interjection by individual customers who are already waiting in the queue. Using matrix-analytic methods and stochastic comparison methods, the waiting times of normal customers and interjecting customers are being studied. The impacts of the two parameters on the waiting times are analyzed in detail, and the implications of the results are discussed with numerical examples.

It is found that the waiting times are sensitive to the tolerance level of interjection by individual customers. It is also found that eliminating customer interjection would be always beneficial to normal customers and arbitrary customers though it would not always be so for interjecting customers.

Chapter 1

Introduction

1.1 Motivation

All of us experience waiting in lines to receive some kind of service in our daily life, and sometimes this waiting time is annoying. Supermarkets, banks, airports and movie theatres are just some of many examples in which we see queues. Production lines, supply chain design, transportation systems (both surface and air traffic) and communications networks are some of the modern fields in which queueing situations are encountered. The customers are sensitive to their waiting times in a queueing system and expect their waiting times to be as small as possible. This fact was a good indication to concentrate this work on analyzing the waiting time of customers. Note that waiting times are among the most important and extensively studied performance measures of queueing theory (Gross and Harris [16], Kleinrock [23], Neuts [27], White [37] and Cohen [9]).

We usually experience times when we are waiting in queue and see that some customers do not join the queue at the end like normal customers, but try to cut in the queue hoping to have a shorter waiting time, and thus a higher level of satisfaction. This behaviour is called "*customer interjection*". For instance, the first-come-first-served (FCFS) service discipline is usually assumed in public places like restaurants, banks, airports and supermarkets. However, customer interjections can still be seen in these places. For traffic at an intersection with several lanes, some drivers may use the right (left) turning lane to cut into a straight lane or vice versa. Such interjections cause the traffic to slow down in order to avoid possible accidents. Since customer interjection can decrease the efficiency of queueing systems, it is an interesting topic for further investigation.

In telecommunications networks, users send data to the network and then the server processes the data packages and transfers them to the final destinations (users). Sometimes, in order to test the efficiency of transmission and performance of the network, artificial (since they are not sent by actual users of the network) data packages are inserted into the normal traffic of the system in a random manner. The goal is to measure and improve the network performance by tracking some marked data inserted into the network. However, it may force extra workload to the system and interrupt the actual performance of the network. Considering the whole network as a queueing system, the data sent by the network users are normal customers of the system, and the data inserted purposely into the system are the interjecting customers. This study is motivated by the design and analysis of the above systems.

1.2 Problem of Interest

The queueing model of interest is a single server queueing system, in which all customers join a single queue and there are two types of customers. The two types of customers are called “normal customers” and “interjecting customers” respectively. When a normal customer arrives, it joins the queue at the end. When an interjecting customer arrives, it may join the queue at any position, depending on the queue length at the arrival epoch and certain given probability distribution. In this work, the focus is on two parameters related to social behaviour and individual tolerance on interjection. The goal is to analyze the impacts of the two parameters on the waiting time of different types of customers, both theoretically and numerically.

1.3 Scope of the Thesis

The theory of Markov processes (Cohen [9], Ross [30]) provides the basic tool to study the queueing systems of interest. In addition, stochastic comparison (Chen and Yao [7], Ross [30], and Shake and Shanthikumar [33]), matrix analytical methods (Neuts [28], [27]) and Laplace transform (Kleinrock [23], White [37], Gross and Harris [16]) are utilized in the study as well. Particularly, the monotonicity of the waiting time in some system parameters is shown using stochastic comparison methods. Algorithms for computing the mean and variance of waiting times are developed based on matrix analytical methods. Waiting time distributions and their moments are found using the Laplace transform technique.

1.4 Organization of the Thesis

In Chapter 2, the existing literature related to the subject of this work is summarized. In Chapter 3, the queueing process of the system of interest is described. In Section 3.2, the performance measures that have been used in the study of the model are introduced. In Section 3.3, some known results for the classical M/M/1 queue with regard to its queue length and the waiting time of customers are stated. Chapter 4 is the theoretical part of the study. There are 4 sections in this chapter. In Section 4.1, the waiting time of a customer currently in position *n* is studied. In Sections 4.2 and 4.3, the waiting time of an arbitrary normal customer and the waiting time of an arbitrary interjecting customer are investigated, respectively. In Section 4.4, the waiting time of an arbitrary customer is investigated. Numerical examples are presented in Chapter 5. An efficient algorithm for computing the mean and variance of the waiting times is presented in this chapter as well. Lastly, Chapter 6 provides the concluding remarks concerning the research conducted for this work. A brief summary of the prominent results and suggestions for possible future research are given. In Appendix A, the stochastic comparison method is discussed briefly. Appendix B provides the proofs for some equations and relationships studied in this work.

Chapter 2

Literature Review

Queueing theory is one of the most important branches of modern probability theory that has been studied extensively since the beginning of the past century. Queueing theory attempts to describe queues through detailed mathematical analysis. As a result of its applications in industries, technology, telecommunications networks, information technology and management sciences, it has been an interesting research area for many researchers active in this field. Kleinrock ([23], [24]) has done extensive work on the theory of queueing systems and their computer applications. Takagi [34] considers queueing phenomena with regard to its applications in performance evaluation of computer and communication systems. Chen and Yao [7] focus on analysing the performance and optimization of queueing networks.

2.1 History

A. K. Erlang, a Danish engineer, was the founder of queueing theory and published "*The theory of probabilities and telephone conversations*" as the first paper on queueing theory in 1909. According to Cohen [9], the development of queueing theory until about 1940 has mainly been directed by the needs encountered in the design of automatic telephone exchanges. He believes that after the World War II, when applications of mathematical models and methods in technology and organizations rose to a level hither to unknown, it was soon recognized that queueing theory had a very broad field of application. Bhat [6] in "Sixty years of queueing theory," has reviewed the developments of queueing theory with respect to its three constituent problems: "1- Behavioural, 2- Statis-tical, and 3-

Operational." He believes that by looking at the general picture of the developments in these areas, it is not hard to conclude that the systems studied have become more and more realistic over the years. He has cited 94 references in his publication which is a good summary over the literature on queueing theory within its first sixty years of existence. In 1969, Satty [31] in "Seven more years of queues: A Lament and a bibliography," stated that, "In the past seven years the literature on queueing theory has increased by half of its amount for the previous fifty years. Improvements do not match the increase in theoretical developments. Rarely has so much ingenuity been shown in tackling a variety of technical problems on paper by some of the ablest people in the world. But real life queues are still primitive."

The first notation for characterising the queueing system was suggested by Kendall [21] in 1953. He introduced a three-factor **A/B/C** notation in which **A** refers to arrival (or interarrival) distribution, **B** refers to the service time distribution and **C** refers to the number of servers in the system. Based on his notation, the most common values of **A** and **B** are as follows:

Symbol	**Explanations**
M	Exponential
D	Deterministic
E_k	Erlang type k (k=1, 2, ...)
H_k	Hyperexponential type k
PH	Phase type
G	General

Cohen [9] and Kleinrock [23] are among researchers who have discussions about M/M/1, M/M/c, G/M/1, G/M/c, M/G/1, G/G/1 and G/G/c queueing systems. Gross and Harris [16] also consider other queueing models like M/E_k /1 and E_k /M/1 in addition to those models.

Gross and Harris [16] have discussed 6 characteristics that in most cases are adequate to describe a queueing system:

I. Arrival pattern of customers
II. Service pattern of servers
III. Queue discipline
IV. System capacity
V. Number of service channels
VI. Number of service stages.

Gross and Harris [16] also discuss the reaction of a customer upon entering the system. "A customer may decide to wait no matter how long the queue becomes, or if the queue is too long to suit him, may decide not to enter it. If a customer decides not to enter the queue upon arrival, he is said to have *balked.* On the other hand, a customer may enter the queue, but after a time lose patience and decide to leave. In this case he is said to have *reneged*. In the event that there are two or more parallel waiting lines, customers may switch from one to another, that is, *jockey* for position. These three situations are all examples of queues with *impatient customers*." This work considers the fourth group of impatient customers, who are the ones that try to cut in the queue to get a faster service and are called *interjecting customers*.

2.2 Queueing Systems with Interjecting Customers

There is limited study on queues with customer interjections in the existing literature. Larson [25] discussed social justice and the psychology of queueing. He used "slips" and "skips" to describe customer interjections: "He who experiences a slip is victimized; he who skips gets a certain sense of satisfaction from his good fortune." A number of examples where slips and skips occur are presented and analyzed in his work and Larson explains how some of them could have significant monetary consequences. For some cases, some nontechnical yet effective solutions are discussed. He also explains that the feedback on for example the estimated amount of time customers have to wait in queue usually makes them feel better.

In Gordon [15], some probability laws for slips and skips in a number of different popular queueing systems were derived. The models considered in his work are different from the model analyzed in this work. In particular, they considered the situation where first-in customers leave the queueing system later than others. The difference between the order of arrival and the order of departure is caused mainly by service, not interjection. In Whitt [38], a similar issue is considered for queueing networks. He refers to slips and skips as "customer overtaken" in his study and says that, "Customer B overtakes customer A in a queueing system if A arrives before B but B departs first."

2.3 Priority Queues

Although the existing queueing models are different from the one considered in this work, some of them are closely related. For example, the priority queues have been investigated extensively (Cohen [9], Takagi [35], Takine and Hasegawa [36], and references therein). Drekic and Stanford [11] describe priority systems as queueing systems in which the intention is to give one group of customers (or possibly several) premium service and the ability to preempt is a key element in that strategy. In priority queues, customers with higher priority skip over the customers with lower priority, which leads to the difference between the waiting times of different types of customers. The queueing process of customers with higher priority is, in general, not affected by that of lower priority customers.

Priority queues can be preemptive (any customer with lower priority can be ejected from the server by a higher priority customer) or non-preemptive (service is never interrupted). Preemptive service disciplines are divided into two categories as well. One is preemptive resume in which the accumulated service times are retained for a low-priority customer after preemption. The other is preemptive repeat in which the service starts from the beginning for a low-priority customer after every ejection from the service and the accumulated service effort is therefore lost when preempted.

There are some negative consequences for customers in both pure preemptive and pure non-preemptive priority queues. In preemptive systems, customers of lower priority might experience a longer time in the system because of the higher priority customers. Also, they might be ejected from the server, upon the arrival of a higher-priority customer, several times before they leave the system. Drekic and Stanford [11] talk about typical situations where one would not want to allow preemptions. One situation includes those computing applications where a lengthy service rendered prior to the point of interruption would be irretrievable. On the other hand, in a non-preemptive priority system, customers with high priority who expect an immediate service might be waiting in the queue for a long time until the lower priority customer with a long service time leaves the server.

Pure preemptive and pure non-preemptive priority systems have been investigated in most research but Zeltyn, Feldman and Wassekrug [40] have investigated the waiting and sojourn times in a multi-server queue with mixed priorities (combination of preempting and non-preempting). They explain that contact centers, health care, and communications networks are some examples of service systems that are relevant to systems with mixed priority discipline. They consider a multi-server queue with n servers and K classes of customers. Customers of the P highest priorities ($P<K$) can preempt customers with lower priorities, ejecting them from service and sending them back into the queue. In general, class i customers have higher priorities than class j customers if $i<j$. Customers of classes 1,..., P are named preempting since they can eject lower-priority customers from service. Customers of classes P+1,..., K are named non-preempting since they can enter service only if an idle server is available. In their system, the arrival processes of all classes are independent Poisson processes and service times are exponentially distributed with the same mean for all classes. They calculate the Laplace-Stieltjes transforms of the waiting times of all classes of customers and from there, they find the moments of any

order for the waiting times. In this work, the same method is used to find the mean and variance of the waiting times.

As discussed earlier, there are negative consequences for customers of a lower priority in preemptive priority systems. Drekic and Stanford [11] say that, “If one were to provide some relief for these customers, one might consider either limiting the frequency with which preemptions occur or intervening on a customer’s behalf if its time in system becomes particularly long.” Several researchers have discussed this problem and tried to propose models to minimize the impact of preemption on lower priority classes. Kleinrock [22] has a model of intervention which is based on the customer’s accumulated time in the system. In his model, the customer’s priority is a linear function of its accumulated time in the system. Adiri and Domb ([2], [3]), Paterok and Ettl [29] and Takagi and Kodera [35] have investigated the idea of “preemption distance” discipline which is based on the distance of two customers in the queue. In this model, preemptions are only allowed if the difference between the indices of the two priority classes in the queue (distance) exceeds a specified threshold. Cho and Un [8] study an M/G/1 queue under a combined preemptive/non-preemptive priority discipline. They propose three discretion rules for preemption based on a parameter of the low priority job: the elapsed service time, the ratio of the elapsed to total service time and the remaining service time. They assume that each priority class has the same discretion rule for preemption. Drekic and Stanford [12] have analyzed a priority queueing model in which an upper limit is given on the number of preemptions a customer experiences. They have also developed a hybrid preemption policy for an M/G/1 priority queue in which interventions to improve the delay performance are based on the customer's accumulated service effort. When the amount of service rendered has crossed a specified threshold, further preemptions are blocked and the remaining service effort is rendered on a nonpreemptive basis. Drekic and Stanford [12] state that the decision to preempt or not, based on the accumulated service effort, was first introduced by Avi-Itzhak et al [5] and Jaiswal [19]. Drekic and Stanford [11] consider the Cho and Un [8] intervention policies in their study and describe them as follows:

- Proportion-based (PB) policy: once a certain proportion α, $0 \leq \alpha \leq 1$, of the service time requirement has been successfully rendered, further preemptions are prevented.
- Front-end time-based (FETB) policy: once T time units of service have been rendered successfully, further preemptions are prevented.
- Tail-end time-based (TETB) policy: once the time remaining to successfully complete service is less than τ time units, further preemptions are prevented.

The difference between this model and that of Cho and Un [8] is that they allow each priority class to employ its own service threshold policy. Drekic and Stanford [11] obtain the Laplace-Stieltjes transform and mean of each class sojourn time for the three threshold types, and for both preempive resume and preemptive repeat service disciplines. By numerical examples, they [11] show that, "It is frequently the case that a good combination of preemptible and nonpreemptible service performs better than both the standard preemptive and nonpreemptive queues."

For a queue with customer interjection, the queueing processes of all types of customers interact by others. Thus, while customer interjection can be viewed as partial priority, its impact on the queueing process of different types of customers in queue is different from that of priority queueing systems.

2.4 Queues with Customer Jockeying

After the shortest queue model with only two servers was proposed and solved by Haight [17] in 1958, queues with customer jockeying (or transferring) (He and Neuts [18], Whitt [39], Zhao and Grassmann ([42], [43]), Disney and Mitchell [13], Elsayed and Bastani [14], Kao and Lin [20], Zhao [41], and Aden, Wessels and Zijm [1]) have been studied extensively as well. For such queueing models, customer "slips" and "skips" can occur since there are multiple queues and jockeying. In some cases, customers can choose which queue to join in order to minimize their waiting times. For example, Zhao and Grassmann [42] consider several buffers with infinite capacity in their model. Each arriving customer joins the shortest waiting queue if it cannot be processed immediately. Zhao and Grassmann [42] define a certain threshold value and when the difference of the waiting packet numbers (queue length) between the longest waiting line and the shortest exceeds the threshold value, the last waiting packet is allowed to move (jockey) to the shortest waiting line.

In queueing systems with customer jockeying, decisions reside within individual customers and there is no interaction between customers. In the model investigated in this work, the customer interjection process is between interjection customers and non-interjection customers. The behaviour of either type of customers has great impact on the queueing process of the other.

2.5 Markov Process

The theory of Markov processes provides the basic tool for the analysis of queueing systems. The following review is a brief summary of existing literature on the theory of Markov processes.

A rather recently developed branch of probability theory is the theory of stochastic processes. Brownian motion is one of the first processes that has been studied extensively since the beginning of the past century and from its investigation, fundamental contributions to probability theory have originated. Gross and Harris [16] say that, "A stochastic process is the mathematical abstraction of an empirical process whose development is governed by probabilistic laws." A stochastic process is usually defined as a family of random variables, $\{X(t), t \in T\}$, in which parameter t stands for time and $X(t)$ denotes the state of the process at time t. If $t = 0,1,2,...,$ then the stochastic process $\{X(t), t \in T\}$ is said to be a discrete-time process, and if $t \in [0, \infty),$ the process has continuous-time parameter. Poisson processes and birth-and-death processes are two well-known stochastic processes that have been studied and used extensively. A continuous-time stochastic process with Markovian property is called a *Markov process*. In other words, a Markov process is a continuous-time stochastic process in which the future condition of the process is just dependent on the current condition of the process and is independent of the past. This property is called the *memoryless property*. Ross [30] has the following mathematical notation for Markov processes: "Consider a continuous-time stochastic process $\{X(t), t \in T\}$ taking on values in the set of nonnegative integers.

We say that this process is a continuous-time Markov chain if for all $\{X(t), t \in T\}$ and nonnegative integers $i, j, x(u), 0 \le u \le s$,

$$P\{X(t+s) = j \mid X(s) = i, X(u) = x(u), 0 \le u < s\} = P\{X(t+s) = j \mid X(s) = i\}. \quad (2.5.1)$$

If, in addition, $P\{X(t+s) = j \mid X(s) = i\}$ is independent of s, then the continuous-time Markov chain is said to have stationary or homogenous transition probabilities." The queueing system studied in this work is a time-homogenous Markov process with stationary transition probabilities. A Markov chain is said to be ergodic if it is irreducible (all states are reachable from all other states), aperiodic (for each state, the probability of returning to that state is positive for all steps), and non-null recurrent (each state is revisited again with probability 1, and the mean time to return to each state is finite). If a Markov chain is ergodic, then its steady-state distribution exists. More comprehensive and detailed discussions about the Markov processes can be found in the works of Cohen [9], Gross and Harris [16], Kleinrock [23], and Ross [30].

2.6 Laplace Transform

Transform techniques (especially the Laplace transform) have made a great contribution to the analysis of queueing systems. Gross and Harris [16] mention that, "A transform is merely a mapping of a function from one space to another." The usefulness of transform techniques comes from the fact that it is easier to solve some problems after the application of a transform on them rather than their initial form. Then the solution of the transformed problem can be inverted to its initial form. White [37] emphasizes that transform methods are basically a convenience; that is, they frequently offer a simplified approach to the solution of systems problems. White [37] mentions that the applications of the Laplace transform in the context of queueing theory are, “Principally to solve certain differential equations, to identify the distribution of continuous random variables, and in some cases to determine the moments of continuous random variables.” The Laplace transform is defined (White [37], Ross [30], Kleinrock [23], Gross and Harris [16]) as follows:

Let X denote a non-negative random variable with probability density function $f(x)$ $x \geq 0$, then the Laplace transform of $f(x)$ is defined by

$$L[f(x)] = \int_0^\infty f(x) e^{-sx} dx = E[e^{-sX}] \qquad (2.6.1)$$

where s is a complex variable given by $s = \sigma + i\omega$, $i = \sqrt{-1}$. Since the LT maps a function of x into a new function of s, it can also be denoted by $\bar{f}(s)$ or $\overline{F}(s)$. In this work, the Laplace transform is used to find the waiting time distributions and also the mean and variance of the waiting times. The first and second moment of the random variable X can be obtained from the Laplace transform as follows:

$$E(X) = -\frac{d\bar{f}(s)}{ds}\Big|_{s=0},$$

$$E(X^2) = \frac{d^2\bar{f}(s)}{ds^2}\Big|_{s=0.} \qquad (2.6.2)$$

If one were to review the existing literature regarding queueing systems, especially works in priority queueing systems and queues with customer jockeying, it would be realized that almost all researchers have used the Laplace transform technique to analyze the waiting time and sojourn times in their queueing system of interest.

Chapter 3

The Model and Basic Analysis

In Section 3.1 of this chapter, the queueing process of the system of interest is described. In Section 3.2, two measures to evaluate the performance of the system are defined. These two measures are the queue length and the waiting time of customers. The waiting times of customers are the main concern of this work. In Section 3.3, some known results of the classical M/M/1 queue with regard to its queue length and the waiting time of customers are stated.

3.1 An M/M/1 Queue with Customer Interjection

A single server queueing system with a single queue and two types of customers are assumed in this work. The two types of customers are called "normal customers" and "interjecting customers" respectively. When a normal customer arrives, it joins the queue at the end. When an interjecting customer arrives, it may join the queue at any position depending on the queue length at the arrival epoch and certain given probability distribution.

It is assumed that customers enter the system according to a Poisson process with arrival rate λ. The service times of all customers are independent and identically distributed random variables (i.i.d.r.v) with an exponential distribution with service rate μ. The service process and the arrival process are independent. It is assumed that an arriving customer is interjecting with probability η_i $(0 \le \eta_i \le 1)$ and is normal with probability $1-\eta_i$. Note that subscript "i" is for the word "interjection". According to a classical result about Poisson process (Ross [30]), such an arrival process can also be viewed as the superposition of two independent Poisson processes with arrival rate $\eta_i\lambda$ and $(1-\eta_i)\lambda$, respectively. When an interjecting customer arrives, it tries to join the queue as close to the head of the queue as possible. It is assumed that the interjecting customers do not interrupt the service in process, if a service is being rendered. The arriving customer contacts the first customer waiting in queue for possible interjection. That customer (regardless of its own type) may let the new customer cut in with probability η_c $(0 \le \eta_c \le 1)$ (i.e., taking the first position in queue). Note that the subscript "c" is for "cutting in". If the first customer refuses the interjection request, the interjecting customer contacts the second customer in queue. The process continues until either the customer interjects successfully or joins the queue at the end if all waiting customers refuse interjection. It is assumed that the time for finding a position in the queue is negligible for an interjecting customer. Given that there are n customers in queue at the arrival epoch, there are positions {1, 2, 3,..., n+1} available to the arriving interjecting customer. Thus, the distribution of the position taken by the interjecting customer has a stopped geometric distribution $\{\eta_c,(1-\eta_c)\eta_c,(1-\eta_c)^2\eta_c,\ldots,(1-\eta_c)^{n-1}\eta_c,$ $(1-\eta_c)^n\}$ on positions {1, 2, 3,..., n+1}. Note that the probability to occupy $(n+1)st$ position by an arriving interjecting customer is equal to the rejection of interjection request from all n customers in the queue. Since the probability of rejecting interjection by each customer is $(1-\eta_c)$, then $(1-\eta_c)^n$ is the probability to take the $(n+1)st$ position by an arriving customer. The parameter η_i represents the percentage of customers with interjection intention, which reflects the societal behaviour on interjection. The parameter η_c represents the level of tolerance of individuals on interjection. From social justice point of view, it is always expected that the values of η_i and η_c should be low, i.e., close to 0.

3.2 Performance Measures: Definition and Notation

After describing the queueing process of the desired queueing model in Section 3.1, interest centers on evaluating the performance of the system of interest. The first step is to define some appropriate performance measures. Two measures of performance are studied in this work, which are queue length and waiting time of customers. By reviewing the existing literature on queueing theory, it would be realized that both of these random variables are standard measures to evaluate the behaviour of a queueing system. Cohen [9], Gross and Harris [16], Ross [30] and Kleinrock [23] have extensively talked about these two performance measures in a number of popular queueing systems.

Queue length at time t is equal to the total number of customers in the system at time t. In order to have a stable system, queue length should be finite. The average queue length in the steady state is investigated in this work.

On the other hand, the main concern of this work is to investigate the impact of customer interjection on the behaviour of an M/M/1 queueing system. Thus, the waiting times of customers are interesting performance measures since it is possible to compare the impact of customer interjection on the waiting time of customers studied in this work with the waiting time of customers in an M/M/1 queueing system with FCFS service discipline. The waiting time of a customer in the queue is equal to the length of time starting from the epoch that a customer enters the queue and ending at the epoch that the customer enters the server. This work is concentrated on analyzing the impact of the pair (η_i, η_c) on the mean and variance of the waiting times of normal customers and interjecting customers, both theoretically and numerically. The notations regarding the performance measures of the model studied in this work are described as follows:

$q(t)$: Queue length at time t (the total number of customers in the system)

$W_n(\eta_i, \eta_c)$: Waiting time of a customer currently in position n

(The length of the time starting from the epoch that a customer is currently in position n in the queue and ending at the epoch that the customer enters the server)

$W(\eta_i, \eta_c)$: Waiting time of an arbitrary normal customer in the queue

$V(\eta_i, \eta_c)$: Waiting time of an arbitrary interjecting customer in the queue

$W_a(\eta_i, \eta_c)$: Waiting time of an arbitrary customer in the queue

3.3 Basic Analysis of Queue Length and Waiting Time

It is readily seen that the queue length in the system of interest is the same as that in the classical M/M/1 queue with a FCFS service discipline, but the waiting time can be different. First, some known results of the classical M/M/1 queue are stated for later use. Let $q(t)$ be the queue length at time t (the total number of customers in the system at time *t*). Then the steady state distribution of $q(t)$ exists if and only if $\rho = \lambda / \mu < 1$, and is given by: (Cohen [9], Ross [30])

$$\pi_n = \lim_{t \to \infty} P\{q(t) = n\} = (1-\rho)\rho^n, n \geq 0; \tag{3.3.1}$$

and the average queue length is equal to

$$\begin{aligned} \lim_{t \to \infty} E\left(q(t) \right) &= \sum_{n=0}^{\infty} n\pi_n = \sum_{n=0}^{\infty} n(1-\rho)\rho^n = (1-\rho)\rho \sum_{n=1}^{\infty} n\rho^{n-1} \\ &= (1-\rho)\rho \frac{d\left(\sum_{n=0}^{\infty} \rho^n \right)}{d\rho} = (1-\rho)\rho \frac{d\left(\frac{1}{1-\rho} \right)}{d\rho} = \frac{(1-\rho)\rho}{(1-\rho)^2} = \frac{\rho}{(1-\rho)}. \end{aligned} \tag{3.3.2}$$

In this section and Sections 4.2 and 4.4, it is assumed that $\rho < 1$. Let $W(\eta_i, \eta_c)$ be the waiting time of a customer in steady state. Then, for $t \geq 0$, the following relationships have already been verified: (Gross and Harris [16], Cohen [9])

$$\begin{cases} P\{W(0,0) \leq t\} = 1 - \rho e^{-(\mu - \lambda)t}; \\ E[W(0,0)] = \dfrac{\rho}{(\mu - \lambda)}; \\ Var(W(0,0)) = \dfrac{\rho(2-\rho)}{(\mu - \lambda)^2}, \end{cases} \tag{3.3.3}$$

where "Var" is for the variance of a random variable. Note that $(\eta_i, \eta_c) = (0,0)$ implies that the system is equivalent to the classical M/M/1 queue. It is intuitive that $W(\eta_i, \eta_c)$ is (stochastically) larger if η_i and η_c are larger. In this work, it will be shown that $W(\eta_i, \eta_c)$ increases in η_i and η_c with respect to the stochastically larger order. Furthermore, it is more interesting to find out at which point of (η_i, η_c), the increase in

$W(\eta_i, \eta_c)$ (and other waiting times) is significant, which is the main subject of this work. To answer such a question, the analysis will concentrate on the mean and variance of $W(\eta_i, \eta_c)$, $V(\eta_i, \eta_c)$ and $W_a(\eta_i, \eta_c)$, respectively.

Chapter 4

Waiting Time Analysis

In this chapter, the focus is on analyzing the waiting time of customers in the queueing system of interest theoretically. The impacts of the pair (η_i, η_c) on the waiting times of a customer in position n, an arbitrary normal customer, an arbitrary interjecting customer and an arbitrary customer are studied respectively. The Laplace transform technique is applied to find the mean and variance of the waiting time of each type of customer.

4.1 Waiting Time of a Customer in Position *n*

Let $W_n(\eta_i, \eta_c)$ be the length of the time starting from the epoch that a customer is currently in position n in the queue and ending at the epoch that the customer enters the server, $n \geq 0$. To find the distribution of $W_n(\eta_i, \eta_c)$, an absorbing Markov process is introduced to describe the change of the position of a customer in the queue. Suppose that a customer is in position n in the queue. The customer moves to position $n-1$ if the current service completes before the next arrival. If the next arrival occurs first, the customer remains in position n if the arrival does not interject. If the new customer interjects into one of the first n positions, the customer in position n moves to position $n+1$. The new customer interjects into one of the first n positions with probability $\eta_i\left(1-(1-\eta_c)^n\right)$. Therefore; the change of the position of a customer in queue can be described by a Markov process with a state space $\{0,1,2,3,...\}$ and an infinitesimal generator that is described as follows:

$$Q_a = \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \end{array} \begin{bmatrix} 0 & 0 & 0 & 0 \\ \mu & -\mu - \lambda\eta_i\eta_c & \lambda\eta_i\eta_c & 0 \\ 0 & \mu & -\mu - \lambda\eta_i(1-(1-\eta_c)^2) & \lambda\eta_i(1-(1-\eta_c)^2) \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} \qquad (4.1.1)$$

$$\equiv \begin{bmatrix} 0 & 0 \\ \mathbf{b} & Q \end{bmatrix},$$

where **b** is a column vector with all elements being zero except the first, which is μ. By definition, $W_n(\eta_i, \eta_c)$ is the absorption time of state zero of the Markov process Q_a, given that the initial state is n , $n \geq 1$. Thus, the distribution of $W_n(\eta_i, \eta_c)$ has a phase type representation (Neuts [27]) with an infinite state space and

$$P\{W_n(\eta_i, \eta_c) < t\} = 1 - \boldsymbol{\alpha}_n \exp\{Qt\}\mathbf{e}, \qquad (4.1.2)$$

where $\boldsymbol{\alpha}_n$ is a vector with the *n*th component being one and all others zero and **e** is the column vector with all components being one. By a well-known ergodicity condition for birth-and-death process (Cohen [9]), the following lemma gives a necessary and sufficient condition for the finiteness of the waiting times.

Lemma 4.1.1 The waiting times $\{W_n(\eta_i, \eta_c), n \geq 0\}$ and their moments are finite if and only if $\eta_i \rho < 1$.

Proof. For a customer currently in position *n*, the rate with which a transition occurs from *n* to *n*+1 is equal to $\lambda\eta_i(1-(1-\eta_c)^n)$, and the rate with which a transition occurs from *n* to *n*-1 is equal to μ. In order to have a stable system in which queue length and the waiting times are finite, the service rate must be bigger than the rate with which a transition occurs from position *n* to position *n*+1, for all *n*, which can be written as:

$$\mu > \lambda\eta_i(1-(1-\eta_c)^n) \qquad \text{for all } n,\ n=0,1,2,\ldots\ .$$

On the other hand, $\lim_{n\to\infty}(1-\eta_c)^n = 0$ and $\lim_{n\to\infty}(1-(1-\eta_c)^n)=1$. Let X_k denote $\lambda\eta_i(1-(1-\eta_c)^k)$, which is the rate with which a transition occurs from state k to k+1. Then it is easy to find that $X_0 < X_1 < X_2 < \ldots$. So the necessary and sufficient condition that $\{W_n(\eta_i,\eta_c), n\ge 0\}$ and their moments are finite is that $\mu > \lambda\eta_i$ or $\rho\eta_i < 1$. This completes the proof of Lemma 4.1.1.

It is assumed that $\eta_i\rho < 1$ in the rest of this work. In Section 4.2 and 4.4, the condition $\eta_i\rho < 1$ is replaced by a stronger one $\rho < 1$. It is intuitive that $W_n(\eta_i,\eta_c)$ increases in η_i and η_c. Based on Lemma 4.1.1 and equations (4.1.1) and (4.1.2), the monotonicity of $W_n(\eta_i,\eta_c)$ in η_i and η_c can be shown (Shaked and Shanthikumar [33]).

Lemma 4.1.2 Assume that $\eta_i\rho < 1$. If $(\eta_i,\eta_c) \le (\eta_i', \eta_c')$, i.e. $\eta_i \le \eta_i'$ and $\eta_c \le \eta_c'$, then the waiting time $W_n(\eta_i',\eta_c')$ is larger than the waiting time $W_n(\eta_i,\eta_c)$ with respect to the stochastically larger order.

Proof. By equation (4.1.2), the following relationship can be obtained:

$$P\{W_n(\eta_i,\eta_c) > t\} = e^{-(\lambda+\mu)t}\boldsymbol{\alpha}_n \exp\{(I + \frac{Q}{(\lambda+\mu)})(\lambda+\mu)t\}\mathbf{e}$$

$$= e^{-(\lambda+\mu)t}\sum_{m=0}^{\infty}\frac{((\lambda+\mu)t)^m}{m!}\boldsymbol{\alpha}_n P^m \mathbf{e}, \tag{4.1.3}$$

where I is the identity matrix and $P = I + \frac{Q}{(\lambda+\mu)} = (p_{i,j})_{i,j\ge 1}$, which is a substochastic matrix (Seneta [32]). To prove the lemma, it is sufficient to show that $P^m\mathbf{e}$ is non-decreasing in η_i and η_c elementwise for $m \ge 0$. It is easy to see that the matrix P is monotone, i.e., the $(k+1)st$ row of P dominates the $(k)th$ row of P for all $k \geq 0, \sum_{j=n}^{\infty} p_{k+1,j} \geq \sum_{j=n}^{\infty} p_{k,j}$, for $n \geq 1$ (Marshal and Olkin [26]). It is easy to obtain that $P\mathbf{e} = (\frac{\lambda}{(\lambda+\mu)}, 1, 1, ...)^T$**,** where "T" is for matrix transpose. Thus the components of $P\mathbf{e}$ are non-decreasing. Denote by $P^m\mathbf{e} = (d_{m,1}, d_{m,2}, ...)^T$. Suppose that $d_{m,1} \leq d_{m,2} \leq \dots$. For $P^{m+1}\mathbf{e}$, it is easy to find that, for $j \geq 2$,

$$
\begin{aligned}
d_{m+1,j} &= p_{j,j-1} d_{m,j-1} + p_{j,j} d_{m,j} + p_{j,j+1} d_{m,j+1} \\
&\geq p_{j,j-1} d_{m,j-2} + p_{j,j} d_{m,j-1} + p_{j,j+1} d_{m,j} \\
&= p_{j-1,j-2} d_{m,j-2} + p_{j-1,j-1} d_{m,j-1} + p_{j-1,j} d_{m,j} \\
&\quad + (p_{j,j} - p_{j-1,j-1}) d_{m,j-1} + (p_{j,j+1} - p_{j-1,j}) d_{m,j} \\
&= d_{m+1,j-1} + \frac{\lambda \eta_i (1-\eta_c)^{j-1} \eta_c}{(\lambda+\mu)} (d_{m,j} - d_{m,j-1}) \geq d_{m+1,j-1}.
\end{aligned} \tag{4.1.4}
$$

By induction, $P^m\mathbf{e}$ is monotone for all m.

Suppose that $(\eta_i, \eta_c) \leq (\eta_i^{'}, \eta^{'}_c)$. Then P is dominated by P' (i.e., every row of P is dominated by corresponding row of P'), which can be readily obtained by comparing the vectors

$$
\left(\frac{\mu}{(\lambda+\mu)}, \frac{1-\mu}{(\lambda+\mu)} - \frac{\lambda \eta_i (1-(1-\eta_c)^n)}{(\lambda+\mu)}, \frac{\lambda \eta_i (1-(1-\eta_c)^n)}{(\lambda+\mu)} \right)
$$

and

$$
\left(\frac{\mu}{(\lambda+\mu)}, \frac{1-\mu}{(\lambda+\mu)} - \frac{\lambda \eta_i' (1-(1-\eta_c')^n)}{(\lambda+\mu)}, \frac{\lambda \eta_i' (1-(1-\eta_c')^n)}{(\lambda+\mu)} \right).
$$

Since $P\mathbf{e}=P'\mathbf{e}$, then $P\mathbf{e}$ is dominated by $P'\mathbf{e}$. Suppose that $P^m\mathbf{e}$ is dominated by $(P')^m\mathbf{e}$. Since P is dominated by P' and $(P')^m\mathbf{e}$ is monotone, it is correct to conclude that $P(P')^m\mathbf{e} \le P'(P')^m\mathbf{e}$. Then $P^{m+1}\mathbf{e} = P(P)^m\mathbf{e} \le P(P')^m\mathbf{e} \le P'(P')^m\mathbf{e} = (P')^{m+1}\mathbf{e}$. Therefore, the elements of $P^m\mathbf{e}$ are monotone in (η_i,η_c) for all m. By equation (4.1.3), the waiting time is monotone in (η_i,η_c) with respect to the stochastically larger order. This completes the proof of Lemma 4.1.2.

Let $w_n^*(s)$ be the Laplace transform of $W_n(\eta_i,\eta_c)$. Denote by $w_n^*(s) = E[\exp\{-sW_n(\eta_i,\eta_c)\}]$, $s \ge 0$. Conditioning on the next transition of the Markov process Q_a, it is easy to obtain that $w_0^*(s)=1$ since $W_0(\eta_i,\eta_c)=0$, and, for $n \ge 1$,

$$w_n^*(s) = \frac{\mu+\lambda\eta_i(1-(1-\eta_c)^n)}{s+\mu+\lambda\eta_i(1-(1-\eta_c)^n)}\left[\frac{\mu w_{n-1}^*(s)}{\mu+\lambda\eta_i(1-(1-\eta_c)^n)}+\frac{\lambda\eta_i(1-(1-\eta_c)^n)w_{n+1}^*(s)}{\mu+\lambda\eta_i(1-(1-\eta_c)^n)}\right]. \quad (4.1.5)$$

Proof. The following independent and identically distributed random variables are defined:

Y_1: The amount of time that a customer currently in position *n* waits until an arriving customer interjects into any of the first *n* positions;

$Y_1 \sim \exp(\lambda\eta_i(1-(1-\eta_c)^n))$;

Y_2: The amount of time that a customer currently in position *n* waits until the customer currently in the server leaves;

$Y_2 \sim \exp(\mu)$;

X_n : The amount of time that a customer currently in position n waits until

a change in her/his position happens;

$X_n \sim \text{Min}\,(Y_1, Y_2) = \exp\ (\mu + \lambda\eta_i(1-(1-\eta_c)^n))$.

Now, it is easy to describe the model and then formulate it. The customer currently in position n waits for a while (X_n) until either an interjecting customer interjects into one of the first n positions available ahead of her/him (birth), or the customer currently in the server leaves (death). If a birth happens, she/he will move into position n+1 and if a death happens, she/he will move into position n-1. Depending on whether a birth or a death happens, her/his waiting time will be equal to $X_n + W_{n+1}(\eta_i, \eta_c)$ or $X_n + W_{n-1}(\eta_i, \eta_c)$ respectively. Now it is easy to formulate the problem by conditioning on the next transition of the Markov process Q_a .

$$
\begin{aligned}
w_n^*(s) &= E[\exp\{-sW_n(\eta_i,\eta_c)\}] \\
&= E[\exp\{-sW_n(\eta_i,\eta_c)\}\,|\,birth]*P(birth) + E[\exp\{-sW_n(\eta_i,\eta_c)\}\,|\,death]*P(death) \\
&= E[\exp\{-s(X_n + W_{n+1}(\eta_i,\eta_c))\}]*\frac{\lambda\eta_i(1-(1-\eta_c)^n}{\mu+\lambda\eta_i(1-(1-\eta_c)^n)} \\
&\quad + E[\exp\{-s(X_n + W_{n-1}(\eta_i,\eta_c))\}]*\frac{\mu}{\mu+\lambda\eta_i(1-(1-\eta_c)^n)} \\
&= \frac{\mu+\lambda\eta_i(1-(1-\eta_c)^n}{s+\mu+\lambda\eta_i(1-(1-\eta_c)^n)}*w_{n+1}^*(s)*\frac{\lambda\eta_i(1-(1-\eta_c)^n}{\mu+\lambda\eta_i(1-(1-\eta_c)^n)} \\
&\quad + \frac{\mu+\lambda\eta_i(1-(1-\eta_c)^n}{s+\mu+\lambda\eta_i(1-(1-\eta_c)^n)}*w_{n-1}^*(s)*\frac{\mu}{\mu+\lambda\eta_i(1-(1-\eta_c)^n)} \\
&= \frac{\mu+\lambda\eta_i(1-(1-\eta_c)^n)}{s+\mu+\lambda\eta_i(1-(1-\eta_c)^n)}\left[\frac{\mu w_{n-1}^*(s)}{\mu+\lambda\eta_i(1-(1-\eta_c)^n)}+\frac{\lambda\eta_i(1-(1-\eta_c)^n)w_{n+1}^*(s)}{\mu+\lambda\eta_i(1-(1-\eta_c)^n)}\right].
\end{aligned}
$$

This completes the proof of equation (4.1.5).

By equation (4.1.5), the following expression for $w_n^*(s)$ can be obtained, which is useful for finding the mean and variance of the waiting times.

Lemma 4.1.3 Assume that $\eta_i \rho < 1$. Functions $\{w_n^*(s), n \geq 1\}$ satisfy the following equation:

$$w_n^*(s) = 1 - \frac{s}{\mu}\left[\sum_{m=1}^{n}\sum_{k=m}^{\infty}\left[(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\right]w_k^*(s)\right]. \tag{4.1.6}$$

Proof. By equation (4.1.5), it is easy to obtain that

$$\begin{aligned}
& w_n^*(s) \\
&= \frac{\mu + \lambda\eta_i(1-(1-\eta_c)^n)}{s + \mu + \lambda\eta_i(1-(1-\eta_c)^n)}\left[\frac{\mu w_{n-1}^*(s)}{\mu + \lambda\eta_i(1-(1-\eta_c)^n)} + \frac{\lambda\eta_i(1-(1-\eta_c)^n)w_{n+1}^*(s)}{\mu + \lambda\eta_i(1-(1-\eta_c)^n)}\right] \\
&= \frac{\mu}{s + \mu + \lambda\eta_i(1-(1-\eta_c)^n)}w_{n-1}^*(s) + \frac{\lambda\eta_i(1-(1-\eta_c)^n)}{s + \mu + \lambda\eta_i(1-(1-\eta_c)^n)}w_{n+1}^*(s), \qquad (4.1.7)
\end{aligned}$$

$$\Rightarrow w_n^*(s) - \frac{\mu}{s + \mu + \lambda\eta_i(1-(1-\eta_c)^n)}w_{n-1}^*(s) = \frac{\lambda\eta_i(1-(1-\eta_c)^n)}{s + \mu + \lambda\eta_i(1-(1-\eta_c)^n)}w_{n+1}^*(s).$$

The following relationships can be obtained by multiplying the sides of equation (4.1.7) by $\dfrac{s + \mu + \lambda\eta_i(1-(1-\eta_c)^n)}{\mu}$:

$$\frac{s + \mu + \lambda\eta_i(1-(1-\eta_c)^n)}{\mu}w_n^*(s) - w_{n-1}^*(s) = \rho\eta_i(1-(1-\eta_c)^n)w_{n+1}^*(s),$$

$$\Rightarrow \frac{s}{\mu}w_n^*(s) + w_n^*(s) + \rho\eta_i(1-(1-\eta_c)^n)w_n^*(s) - w_{n-1}^*(s) = \rho\eta_i(1-(1-\eta_c)^n)w_{n+1}^*(s),$$

$$\Rightarrow w_n^*(s) - w_{n-1}^*(s) = -\frac{s}{\mu}w_n^*(s) + \rho\eta_i(1-(1-\eta_c)^n)\left(w_{n+1}^*(s) - w_n^*(s)\right). \tag{4.1.8}$$

For $m \geq 1$,

$$w_m^*(s) - w_{m-1}^*(s) = -\frac{s}{\mu} w_m^*(s) + \rho\eta_i(1-(1-\eta_c)^m)(w_{m+1}^*(s) - w_m^*(s)), \tag{4.1.9}$$

that leads to

$$w_m^*(s) - w_{m-1}^*(s) = -\frac{s}{\mu}\left[\sum_{k=m}^{\infty}\left[(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\right]w_k^*(s)\right]. \tag{4.1.10}$$

Equation (4.1.6) is obtained from equation (4.1.10) according to the following relationships:

$$\begin{aligned}
w_m^*(s) - w_{m-1}^*(s) &= -\frac{s}{\mu}\left[\sum_{k=m}^{\infty}\left[(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\right]w_k^*(s)\right],\\
w_{m-1}^*(s) - w_{m-2}^*(s) &= -\frac{s}{\mu}\left[\sum_{k=m-1}^{\infty}\left[(\rho\eta_i)^{k-m-1}\prod_{j=0}^{k-m-2}(1-(1-\eta_c)^{m+j-1})\right]w_k^*(s)\right],\\
&\vdots\\
w_2^*(s) - w_1^*(s) &= -\frac{s}{\mu}\left[\sum_{k=2}^{\infty}\left[(\rho\eta_i)^{k-2}\prod_{j=0}^{k-3}(1-(1-\eta_c)^{2+j})\right]w_k^*(s)\right],\\
w_1^*(s) - w_0^*(s) &= -\frac{s}{\mu}\left[\sum_{k=1}^{\infty}\left[(\rho\eta_i)^{k-1}\prod_{j=0}^{k-2}(1-(1-\eta_c)^{1+j})\right]w_k^*(s)\right].
\end{aligned} \tag{4.1.11}$$

Now, by adding the right hand sides and the left hand sides of equation (4.1.11), the following result is obtained:

$$w_n^*(s) = 1 - \frac{s}{\mu}\left[\sum_{m=1}^{n}\sum_{k=m}^{\infty}\left[(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\right]w_k^*(s)\right].$$

Note that $\prod_{j=0}^{-1}(\ldots) = 1$ by convention and $w_0^*(s) = 1$. This completes the proof of Lemma 4.1.3.

With the expression in equation (4.1.6), it is possible to derive formulas for the mean and variance of the waiting time $W_n(\eta_i, \eta_c)$. The results are summarized in the following corollary.

Corollary 4.1.4 Assume that $\eta_i \rho < 1$. The first two moments of $W_n(\eta_i, \eta_c)$ have the following expressions:

$$E[W_0(\eta_i, \eta_c)] = 0;$$

$$E[W_0^2(\eta_i, \eta_c)] = 0;$$

and for $n \geq 1$,

$$\begin{cases} E[W_n(\eta_i,\eta_c)] = \dfrac{1}{\mu}\displaystyle\sum_{m=1}^{n}\sum_{k=m}^{\infty}(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j}); \\ E[W_n^2(\eta_i,\eta_c)] = \dfrac{2}{\mu}\displaystyle\sum_{m=1}^{n}\sum_{k=m}^{\infty}(\rho\eta_i)^{k-m}\Big(\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\Big)E[W_k(\eta_i,\eta_c)]; \\ Var[W_n(\eta_i,\eta_c)] = E[W_n^2(\eta_i,\eta_c)] - (E[W_n(\eta_i,\eta_c)])^2. \end{cases} \tag{4.1.12}$$

The proof of Corollary 4.1.4 is presented in Appendix B. The first two moments of $W_n(\eta_i, \eta_c)$ are non-decreasing in η_i and η_c.

Using the expression in equation (4.1.12), the mean and variance of $W_n(\eta_i, \eta_c)$ can be calculated. However, a computational method based on equation (4.1.12) can be numerically unstable and time consuming. A numerically more efficient algorithm for computing the first two moments of $W_n(\eta_i, \eta_c)$ is developed in Chapter 5.

4.2 Waiting Time of an Arbitrary Normal Customer

In this section, the waiting time of an arbitrary normal customer is considered. Recall that a normal customer always joins the queue at the end. Assuming $\rho < 1$ and conditioning on the number of costumers in the system at the arrival epoch, by equation (3.3.1), it is easy to obtain that

$$P\{W(\eta_i, \eta_c) < t\} = \sum_{n=0}^{\infty} (1-\rho)\rho^n P\{W_n(\eta_i, \eta_c) < t\} ;$$

$$w^*(s) = E[\exp\{-sW(\eta_i, \eta_c)\}] = \sum_{n=0}^{\infty} (1-\rho)\rho^n w_n^*(s), s \geq 0 . \tag{4.2.1}$$

By Lemma 4.1.1 and equation (4.2.1), it can be concluded that $W(\eta_i, \eta_c)$ is non-decreasing in η_i and η_c with respect to the stochastically larger order, which is consistent with intuition. Consequently, the first two moments of $W(\eta_i, \eta_c)$ are non-decreasing functions in η_i and η_c. By equations (4.1.12) and (4.2.1), the first two moments of $W(\eta_i, \eta_c)$ are given as:

$$E[W(\eta_i, \eta_c)] = \frac{\rho}{\mu - \lambda} + \frac{(1-\rho)}{\mu} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \rho^{n+k} \eta_i^k \left[\sum_{m=1}^{n} \prod_{j=0}^{k-1} (1-(1-\eta_c)^{m+j}) \right];$$

$$E[(W(\eta_i, \eta_c))^2] = \sum_{n=0}^{\infty} (1-\rho)\rho^n E[W_n(\eta_i, \eta_c))^2]. \tag{4.2.2}$$

The proof of equation (4.2.2) is presented in Appendix B.

It is easy to see from equation (4.2.2) that $E[W(\eta_i, \eta_c)]$ is an increasing convex function of η_i. By using the computational method developed in Chapter 5, the mean and variance of the waiting time of $W(\eta_i,\eta_c)$ can be computed. The following corollaries give some explicit results at the boundary points.

Corollary 4.2.1 Assume that $\rho<1$. The mean and variance of $W(\eta_i,\eta_c)$ at some boundary points are given as follows:

$$\begin{cases} E[W(\eta_i,0)] = \dfrac{\rho}{\mu-\lambda}; \\ E[W(\eta_i,1)] = \dfrac{\rho}{\mu-\lambda}\left(\dfrac{1}{1-\rho\eta_i}\right); \\ E[W(0,\eta_c)] = \dfrac{\rho}{\mu-\lambda}; \\ E[W(1,\eta_c)] = \dfrac{\rho}{\mu-\lambda} + \dfrac{1-\rho}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right]; \\ Var[W(\eta_i,0)] = Var[W(0,\eta_c)] = \dfrac{\rho(2-\rho)}{(\mu-\lambda)^2}; \\ Var[W(\eta_i,1)] = \dfrac{\rho(2-\rho-\rho^2\eta_i)}{(\mu-\lambda)^2(1-\rho\eta_i)^3}. \end{cases} \tag{4.2.3}$$

The proof for Corollary 4.2.1 is presented in Appendix B. Comparing the mean and variance at these boundary points, it is easy to see that

$$\begin{cases} E[W(\eta_i,0)] = E[W(\eta_i,1)](1-\rho\eta_i) < E[W(\eta_i,1)]; \\ Var[W(\eta_i,0)] = Var[W(\eta_i,1)](1-\rho\eta_i)^3/(1-\rho^2\eta_i/(2-\rho)) < Var[W(\eta_i,1)]; \\ E[W(0,\eta_c)] < E[W(1,\eta_c)]. \end{cases}$$

Thus, when customer interjections are tolerated by individual customers (i.e., $\eta_c=1$, regardless of their type), a normal customer's waiting time and its variation can be significantly larger. Similar results can be obtained for the variances at $(0,\eta_c)$ and $(1,\eta_c)$, but the expressions are tedious. Thus, the details are omitted.

Next, a sensitivity analysis is conducted on the mean waiting time $w(\eta_i,\eta_c)=E[W(\eta_i,\eta_c)]$, to know how fast the waiting time changes if η_i and η_c change their values. By routine calculations, equation (4.2.2) leads to

$$\frac{\partial w(\eta_i,\eta_c)}{\partial \eta_i}=\frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}k\rho^{n+k}\eta_i^{k-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right]; \tag{4.2.4}$$

$$\frac{\partial w(\eta_i,\eta_c)}{\partial \eta_c}=\frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^{k}\left[\sum_{m=1}^{n}\sum_{i=0}^{k-1}(m+i)(1-\eta_c)^{m+i-1}\prod_{j=0;\, j\neq i}^{k-1}(1-(1-\eta_c)^{m+j})\right].$$

Corollary 4.2.2 Assume that $\rho<1$. For boundary points $(\eta_i,0)$, $(\eta_i,1)$, $(0,\eta_c)$ and $(1,\eta_c)$, the following equations are given:

$$\begin{cases} \left.\dfrac{\partial w(\eta_i,\eta_c)}{\partial \eta_i}\right|_{\eta_i=0}=\dfrac{\rho^2\eta_c}{\mu(1-\rho)(1-\rho(1-\eta_c))}; \\ \left.\dfrac{\partial w(\eta_i,\eta_c)}{\partial \eta_i}\right|_{\eta_i=1}=\dfrac{(1-\rho)}{\mu}\displaystyle\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}k\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j}); \end{cases} \tag{4.2.5}$$

and

$$\begin{cases} \left.\dfrac{\partial w(\eta_i,\eta_c)}{\partial \eta_c}\right|_{\eta_c=0}=\dfrac{\rho^2\eta_i}{\mu(1-\rho)^2}; \\ \left.\dfrac{\partial w(\eta_i,\eta_c)}{\partial \eta_c}\right|_{\eta_c=1}=\dfrac{\rho^2\eta_i}{\mu(1-\rho\eta_i)}. \end{cases} \tag{4.2.6}$$

By equations (4.2.5) and (4.2.6), it is easy to observe that the derivative of $w(\eta_i,\eta_c)$ is larger, sometimes significantly, at $\eta_i = 1$ ($\eta_c = 0$) than at $\eta_i = 0$ ($\eta_c = 1$). Note that the expression of $\frac{\partial w(\eta_i,\eta_c)}{\partial \eta_c}\Big|_{\eta_c=0}$ has a denominator with the term $(1-\rho)^2$. Thus, for the waiting time of normal customers, it is not that important that there are some people with the intention to interject. It is more important to make sure that no one tolerates interjection, i.e., to keep $\eta_c = 0$, since $w(\eta_i,\eta_c)$ increases significantly at $\eta_c = 0$.

4.3 Waiting Time of an Arbitrary Interjecting Customer

If an interjecting customer arrives and sees n customers waiting in queue, the customer takes position j with probability $(1-\eta_c)^{j-1}\eta_c, 1 \le j \le n$, and position $n+1$ with probability $(1-\eta_c)^n$. Denote by $V(\eta_i,\eta_c)$ the waiting time of an interjecting customer. Assuming $\eta_i \rho < 1$ and conditioning on the number of customers (which is shown with # symbol in this work) in the system at the arrival epoch, the Laplace transform of the waiting time can be obtained as follows:

$$
\begin{aligned}
v^*(s) &= E[\exp\{-sV(\eta_i,\eta_c)\}] = \sum_{n=0}^{\infty} E[\exp\{-sV(\eta_i,\eta_c)\} | \# = n] P(\# = n) \\
&= \sum_{n=0}^{\infty} (1-\rho)\rho^n E[\exp\{-sV(\eta_i,\eta_c)\} | \# = n] \\
&= 1-\rho + \sum_{n=1}^{\infty} (1-\rho)\rho^n E[\exp\{-sV(\eta_i,\eta_c)\} | \# = n] \\
&= 1-\rho + \sum_{n=1}^{\infty} (1-\rho)\rho^n \left[\sum_{j=1}^{n-1} (1-\eta_c)^{j-1}\eta_c w_j^*(s) + (1-\eta_c)^{n-1} w_n^*(s) \right].
\end{aligned} \tag{4.3.1}
$$

There are two conditions that need to be considered in order to find the mean and variance of the waiting time from equation (4.3.1). These two conditions are $\rho < 1$ and $\rho \geq 1, \eta_i \rho < 1$.

If $\rho < 1$, it is easy from equation (4.3.1) to obtain that

$$v^*(s) = 1-\rho+\sum_{n=1}^{\infty}(1-\rho)\rho^n\left[\sum_{j=1}^{n-1}(1-\eta_c)^{j-1}\eta_c w_j^*(s)+(1-\eta_c)^{n-1}w_n^*(s)\right]$$

$$=1-\rho+(1-\rho)\rho w_1^*(s)+\sum_{n=2}^{\infty}(1-\rho)\rho^n\left[\sum_{j=1}^{n-1}(1-\eta_c)^{j-1}\eta_c w_j^*(s)+(1-\eta_c)^{n-1}w_n^*(s)\right]$$

$$=1-\rho+(1-\rho)\rho w_1^*(s)+\sum_{n=2}^{\infty}(1-\rho)\rho^n\sum_{j=1}^{n-1}(1-\eta_c)^{j-1}\eta_c w_j^*(s)+\sum_{n=2}^{\infty}(1-\rho)\rho^n(1-\eta_c)^{n-1}w_n^*(s)$$

$$=1-\rho+(1-\rho)\rho w_1^*(s)+\sum_{j=1}^{\infty}(1-\rho)\rho^n\sum_{n=j+1}^{\infty}(1-\eta_c)^{j-1}\eta_c w_j^*(s)+\sum_{n=2}^{\infty}(1-\rho)\rho^n(1-\eta_c)^{n-1}w_n^*(s)$$

$$=1-\rho+(1-\rho)\rho w_1^*(s)+(1-\rho)\eta_c\sum_{j=1}^{\infty}(1-\eta_c)^{j-1}w_j^*(s)\sum_{n=j+1}^{\infty}\rho^n+\sum_{n=2}^{\infty}(1-\rho)\rho^n(1-\eta_c)^{n-1}w_n^*(s)$$

$$=1-\rho+(1-\rho)\rho w_1^*(s)+(1-\rho)\eta_c\sum_{j=1}^{\infty}(1-\eta_c)^{j-1}w_j^*(s)\frac{\rho^{j+1}}{(1-\rho)}+\sum_{n=2}^{\infty}(1-\rho)\rho^n(1-\eta_c)^{n-1}w_n^*(s)$$

$$=1-\rho+(1-\rho)\rho w_1^*(s)+\rho\eta_c\sum_{j=1}^{\infty}(1-\eta_c)^{j-1}w_j^*(s)\rho^j+\sum_{n=2}^{\infty}(1-\rho)\rho^n(1-\eta_c)^{n-1}w_n^*(s)$$

$$=1-\rho+(1-\rho)\rho w_1^*(s)+\sum_{n=2}^{\infty}(1-\rho)\rho^n(1-\eta_c)^{n-1}w_n^*(s)+\rho\eta_c\sum_{j=1}^{\infty}\rho^j(1-\eta_c)^{j-1}w_j^*(s)$$

$$=1-\rho+\sum_{n=1}^{\infty}(1-\rho)\rho^n(1-\eta_c)^{n-1}w_n^*(s)+\rho\eta_c\sum_{j=1}^{\infty}\rho^j(1-\eta_c)^{j-1}w_j^*(s)$$

$$= 1-\rho+(1-\rho)\sum_{n=1}^{\infty}\rho^{n}(1-\eta_c)^{n-1}w_n^*(s)+\rho\eta_c\sum_{n=1}^{\infty}\rho^{n}(1-\eta_c)^{n-1}w_n^*(s)$$

$$= 1-\rho+(1-\rho(1-\eta_c))\sum_{n=1}^{\infty}\rho^{n}(1-\eta_c)^{n-1}w_n^*(s). \qquad (4.3.2)$$

If $(\rho \geq 1, \eta_i\rho < 1)$, the (total) queue length is infinite and an interjecting customer takes position j in the queue with probability $(1-\eta_c)^{j-1}\eta_c,\ 1 \leq j \leq \infty$. Conditioning on the position an interjecting customer takes in the queue, it is easy to verify that

$$v^*(s) = E[\exp\{-sV(\eta_i,\eta_c)\}] = \sum_{n=1}^{\infty}(1-\eta_c)^{n-1}\eta_c w_n^*(s), \qquad \rho \geq 1, \eta_i\rho < 1. \qquad (4.3.3)$$

The results can be summarized as follows:

$$v^*(s) = E[\exp\{-sV(\eta_i,\eta_c)\}]$$

$$= \begin{cases} 1-\rho+(1-\rho(1-\eta_c))\sum_{n=1}^{\infty}\rho^{n}(1-\eta_c)^{n-1}w_n^*(s), & \rho < 1; \\ \\ \sum_{n=1}^{\infty}(1-\eta_c)^{n-1}\eta_c w_n^*(s), & \rho \geq 1, \eta_i\rho < 1. \end{cases} \qquad (4.3.4)$$

By Lemma 4.1.2, $V(\eta_i,\eta_c)$ is non-decreasing in η_i with respect to the stochastically larger order, which is intuitive. However, $V(\eta_i,\eta_c)$ may not be monotone in η_c. By Corollary 4.1.4 and equation (4.3.2), the first two moments of $V(\eta_i,\eta_c)$ can be obtained as follows.

Corollary 4.3.1 If $\eta_i \rho < 1$, then

$E[V(\eta_i, \eta_c)]$

$$= \begin{cases} \dfrac{\rho}{\mu - \lambda(1-\eta_c)} + \dfrac{(1-\rho(1-\eta_c))}{\mu} \displaystyle\sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \rho^{n+k}\eta_i^k (1-\eta_c)^{n-1} \left[\sum_{m=1}^{n} \prod_{j=0}^{k-1} (1-(1-\eta_c)^{m+j}) \right], \rho < 1; \\ \\ \dfrac{1}{\eta_c \mu} + \dfrac{\eta_c}{\mu} \displaystyle\sum_{n=1}^{\infty}\sum_{k=1}^{\infty} (\rho\eta_i)^k (1-\eta_c)^{n-1} \left[\sum_{m=1}^{n} \prod_{j=0}^{k-1} (1-(1-\eta_c)^{m+j}) \right], \rho \geq 1. \end{cases} \tag{4.3.5}$$

$$E[(V(\eta_i, \eta_c))^2] = \begin{cases} (1-\rho(1-\eta_c)) \displaystyle\sum_{n=1}^{\infty} \rho^n (1-\eta_c)^{n-1} E[(W_n(\eta_i, \eta_c))^2], \rho < 1; \\ \\ \displaystyle\sum_{n=1}^{\infty} \eta_c (1-\eta_c)^{n-1} E[(W_n(\eta_i, \eta_c))^2], \rho \geq 1. \end{cases} \tag{4.3.6}$$

The proof of Corollary 4.3.1 is presented in Appendix B. At some boundary points, explicit results can be obtained for the mean waiting time:

$$\lim_{\eta_c \to 0} E[V(\eta_i, \eta_c)] = \begin{cases} \dfrac{\rho}{\mu - \lambda}, & \rho < 1; \\ \\ \infty, & \rho \geq 1, \rho\eta_i < 1. \end{cases} \tag{4.3.7}$$

$$\lim_{\eta_i \to 0} E[V(\eta_i, \eta_c)] = \begin{cases} \dfrac{\rho}{\mu - \lambda(1-\eta_c)}, & \rho < 1; \\ \\ \dfrac{1}{\eta_c \mu}, & \rho \geq 1, \rho\eta_i < 1. \end{cases} \tag{4.3.8}$$

$$\lim_{\eta_c \to 1} E[V(\eta_i,\eta_c)] = \begin{cases} \dfrac{\rho}{\mu-\lambda\eta_i}, & \rho<1; \\ \\ \dfrac{1}{\mu-\lambda\eta_i}, & \rho \ge 1, \rho\eta_i < 1. \end{cases} \tag{4.3.9}$$

$$\lim_{\eta_i \to 1} E[V(\eta_i,\eta_c)]$$

$$= \begin{cases} \dfrac{\rho}{\mu-\lambda(1-\eta_c)} + \dfrac{(1-\rho(1-\eta_c))}{\mu} \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \rho^{n+k}(1-\eta_c)^{n-1}[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})], \rho<1; \\ \\ \dfrac{1}{\eta_c\mu} + \dfrac{\eta_c}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{k}(1-\eta_c)^{n-1}[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})], \qquad \rho\ge 1, \rho\eta_i<1, \eta_c>0. \end{cases} \tag{4.3.10}$$

Equations (4.3.7) and (4.3.9) show that the mean of the waiting time is smaller at $\eta_c = 1$ than that at $\eta_c = 0$. That fact can be shown by partial deriving the derivatives of $E[V(\eta_i,\eta_c)]$ at $(\eta_i,\eta_c) = (\eta_i,0)$ and $(\eta_i,1)$:

$$\text{If } \rho<1: \quad \begin{aligned} &\frac{\partial E[V(\eta_i,\eta_c)]}{\partial \eta_c}\Big|_{\eta_c=0} = -\frac{\rho^2(1-\eta_i)}{\mu(1-\rho)^2} < 0; \\ &\frac{\partial E[V(\eta_i,\eta_c)]}{\partial \eta_c}\Big|_{\eta_c=1} = -\frac{\rho^2(1-\eta_i)}{\mu(1-\eta_i\rho)} < 0. \end{aligned} \tag{4.3.11}$$

$$\text{If } \rho\ge 1: \quad \begin{aligned} &\frac{\partial E[V(\eta_i,\eta_c)]}{\partial \eta_c}\Big|_{\eta_c=0} = -\infty; \\ &\frac{\partial E[V(\eta_i,\eta_c)]}{\partial \eta_c}\Big|_{\eta_c=1} = -\frac{1}{\mu} < 0. \end{aligned} \tag{4.3.12}$$

From equations (4.3.11) and (4.3.12), it is easy to see that the mean of the waiting time is more sensitive at $\eta_c = 0$ than at $\eta_c = 1$. Notice that when $\rho < 1$ and $\eta_c = 0$, the denominator contains the term $(1-\rho)^2$. In other words, if η_c increases a little bit at $\eta_c = 0$, the mean of the waiting time will decrease significantly; but if η_c decreases a little bit at $\eta_c = 1$, the mean of the waiting time will increase, but not significantly. Thus the mean waiting time of an arbitrary interjecting customers is minimized at $\eta_c = 1$. Details of equations (4.3.11) and (4.3.12) are presented in Appendix B.

4.4 Waiting Time of an Arbitrary Customer

For an arbitrary customer, it is a normal customer with probability η_i and an interjecting customer with probability $1-\eta_i$. Denote by $W_a(\eta_i, \eta_c)$ the waiting time of an arbitrary customer. If $\rho < 1$, by conditioning on the type of the arbitrary arrival, it is easy to find that

$$P\{W_a(\eta_i,\eta_c) < t\} = \eta_i P\{V(\eta_i,\eta_c) < t\} + (1-\eta_i)P\{W(\eta_i,\eta_c) < t\};$$

$$E[W_a(\eta_i,\eta_c)] = \eta_i E[V(\eta_i,\eta_c)] + (1-\eta_i)E[W(\eta_i,\eta_c)]. \tag{4.4.1}$$

Intuitively, since the mean queue length in the queue is independent of the (η_i, η_c), by Little's law (Gross and Harris [16]), the mean waiting time is independent of (η_i, η_c). Therefore,

$$E[W_a(\eta_i,\eta_c)] = \eta_i E[V(\eta_i,\eta_c) + (1-\eta_i)E[W(\eta_i,\eta_c)] = E[W(0,0)] = \frac{\rho}{(\mu-\lambda)}. \tag{4.4.2}$$

Thus, it is only necessary to discuss the variance of $W_a(\eta_i,\eta_c)$. For the variance of an arbitrary customer, it is easy to find that

$$\begin{aligned} Var(W_a(\eta_i,\eta_c)) = (1-\eta_i)Var(W(\eta_i,\eta_c)) + \eta_i Var(V(\eta_i,\eta_c)) \\ + \eta_i(1-\eta_i)(E[W(\eta_i,\eta_c)] - E[V(\eta_i,\eta_c)])^2. \end{aligned} \tag{4.4.3}$$

The proof of equation (4.4.3) is presented in Appendix B.

It is clear from equation (4.4.3) that the waiting time of an arbitrary customer has the smallest variance if $\eta_i = \eta_c = 0$, which is equal to the variance of the waiting time of a customer in an M/M/1 queue with FCFS service discipline.

Chapter 5

Numerical Examples

In order to compute the mean and variance of the waiting time for the different types of customers that have been studied extensively in the previous chapter, a numerically more efficient algorithm for computing the first two moments of $W_n(\eta_i,\eta_c)$ is developed in this chapter. Having the first two moments of $W_n(\eta_i,\eta_c)$ for n=0, 1, 2, 3, ..., it is easy to compute the mean and variance of the waiting time of normal customers, interjecting customers and an arbitrary customer using equations (4.2.1), (4.3.4), (4.4.2) and (4.4.3). The numerical examples are presented in this chapter using this method. In these examples, for different amounts of (η_i,η_c), the mean and variance of the waiting time for all types of customers are computed. The purpose is to analyze the impact of the pair (η_i,η_c) on the waiting time of customers.

5.1 Outline of a Computational Method

Based on the first passage time, a stable algorithm for computing the mean and variance of $W_n(\eta_i,\eta_c)$ is developed. Denote by T_n the first passage time for a customer in position n to move to position $n-1$ in queue. Let

$$Q_{a,n}=\begin{bmatrix} 0 & 0 & 0 & 0 \\ \mu & -\mu-\lambda\eta_i(1-(1-\eta_c)^n) & \lambda\eta_i(1-(1-\eta_c)^n) & 0 \\ 0 & \mu & -\mu-\lambda\eta_i(1-(1-\eta_c)^{n+1}) & \lambda\eta_i(1-(1-\eta_c)^{n+1}) \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}, \quad (5.1.1)$$

with a state space $\{n-1, n, n+1, ...\}$. Denote by $Q_{a,\infty} = \lim_{n\to\infty} Q_{a,n}$ and $T_\infty = \lim_{n\to\infty} T_n$. It is easy to see that T_n is the absorption time of the Markov process $Q_{a,n}$ for $n \le \infty$, given that the Markov process is in state n initially. Similar to the proof of Lemma 4.1.1, it can be shown that T_n is stochastically larger in n (also see Chen and Yao [7]). Thus, $E[T_n]$ and $E[T^2{}_n]$ increase in n. The variable T_∞ can be considered as the length of the busy period of an M/M/1 queue with arrival rate $\lambda\eta_i$ and service rate μ. It is readily seen that

$$
\begin{aligned}
E[e^{-sT_n}] &= \frac{\mu + \lambda\eta_i(1-(1-\eta_c)^n)E[e^{-sT_{n+1}}]E[e^{-sT_n}]}{s+\mu+\lambda\eta_i(1-(1-\eta_c)^n)}, n \ge 1; \\
E[e^{-sT_\infty}] &= \frac{\mu + \lambda\eta_i(E[e^{-sT_\infty}])^2}{s+\mu+\lambda\eta_i}.
\end{aligned}
\tag{5.1.2}
$$

$E[T_\infty]$ and $E[T^2{}_\infty]$ are finite and are given by:

$$
\begin{aligned}
E[T_\infty] &= \frac{1}{\mu - \lambda\eta_i}, \\
E[T^2{}_\infty] &= \frac{2E[T_\infty](1+\lambda\eta_i E[T_\infty])}{\mu-\lambda\eta_i} = \frac{2\mu}{(\mu-\lambda\eta_i)^3}.
\end{aligned}
\tag{5.1.3}
$$

The following relationships of $\{T_n, n \ge 0\}$ can be shown routinely:

$$
\begin{aligned}
E[T_n] &= \frac{1}{\mu} + \frac{\lambda\eta_i(1-(1-\eta_c)^n)}{\mu}E[T_{n+1}]; \\
E[T^2{}_n] &= \frac{2E[T_n] + \lambda\eta_i(1-(1-\eta_c)^n)(E[T^2{}_{n+1}] + 2E[T_n]E[T_{n+1}])}{\mu}.
\end{aligned}
\tag{5.1.4}
$$

Since $W_n(\eta_i, \eta_c) = T_1 + T_2 + ... + T_n$, then it is easy to find that

$$E[W_n(\eta_i,\eta_c)] = E[T_1] + E[T_2] + \ldots + E[T_n] = E[W_{n-1}(\eta_i,\eta_c)] + E[T_n];$$

$$E[W^2{}_n(\eta_i,\eta_c)] = \sum_{j=1}^{n} E[T_j^2] + 2\sum_{1\le i<j\le n} E[T_i]E[T_j] \tag{5.1.5}$$

$$= E[W^2{}_{n-1}(\eta_i,\eta_c)] + E[T_n^2] + 2E[W_{n-1}(\eta_i,\eta_c)]E[T_n].$$

To compute the mean and variance of W_n approximately, a large enough N should be chosen and the following equations should be set: $E[T_N] = E[T_\infty]$ and $E[T_N^2] = E[T_\infty^2]$. Then $E[W_n]$ and $E[W_n^2]$ can be computed by using formulas in equations (5.1.3), (5.1.4) and (5.1.5) for $n \le N$.

Note that, by the monotonicity property of $\{T_n, n \ge 0\}$, it can be obtained that $E[W_N] \le nE[T_\infty]$ and $E[W_N^2] \le nE[T_\infty^2] + n(n-1)(E[T_\infty])^2$. Thus, N can be chosen large enough so that the error in computing the first two moments of waiting times such as $E[W_n(\eta_i,\eta_c)]$ and $E[W_n{}^2(\eta_i,\eta_c)]$ be smaller than any given positive number.

5.2 The Structure of Numerical Examples

By using the computational method developed in Section 5.1, the first two moments of $W_n(\eta_i,\eta_c)$ for n=0, 1, 2, 3, ..., N are computed and based on that, the mean and variance of the waiting time for each type of customer are obtained. The mean and variance of the waiting time of a normal customer are computed using equation (4.2.1). The mean and variance of the waiting time of an arbitrary interjecting customer are computed using equation (4.3.1). The mean and variance of the waiting time of an arbitrary customer are computed using equations (4.4.2) and (4.4.3). Computational experiments are implemented using the "C" computer programming language. The input parameters to the computer programming are λ, μ, N, η_i and η_c. The outputs of the computer programming are the mean and variance of the waiting time of normal customers, interjecting customers and an arbitrary customer, respectively. Two different amounts for ρ have been chosen to implement the computational experiments. The first one is $\rho = 0.2$ which is close to zero and the second one is $\rho = 0.8$ which is close to one.

5.3 Examples for the Waiting Time of an Arbitrary Normal Customer

Example 1. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $W(\eta_i, \eta_c)$ at boundary points are computed in Table 5.1.

Table 5.1 The mean and variance of $W(\eta_i, \eta_c)$ at boundary points when $\rho = 0.2$

λ	μ	ρ	η_i	η_c	**E(*W*)**	**Var(*W*)**
2	10	0.2	0	0	0.025	0.005625
2	10	0.2	0	1	0.025	0.005625
2	10	0.2	1	0	0.025	0.005625
2	10	0.2	1	1	0.03125	0.010742

The results in Tables 5.1 for the mean and variance of $W(\eta_i, \eta_c)$ at boundary points correspond with the results that could be obtained from Corollary 4.1.1 for boundary points when $\rho = 0.2$. This verifies the accuracy of the computational method in computing the mean and variance of the waiting time of normal customers.

Example 2. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $W(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_i \le 1$ are plotted in Figures 5.1 and 5.2.

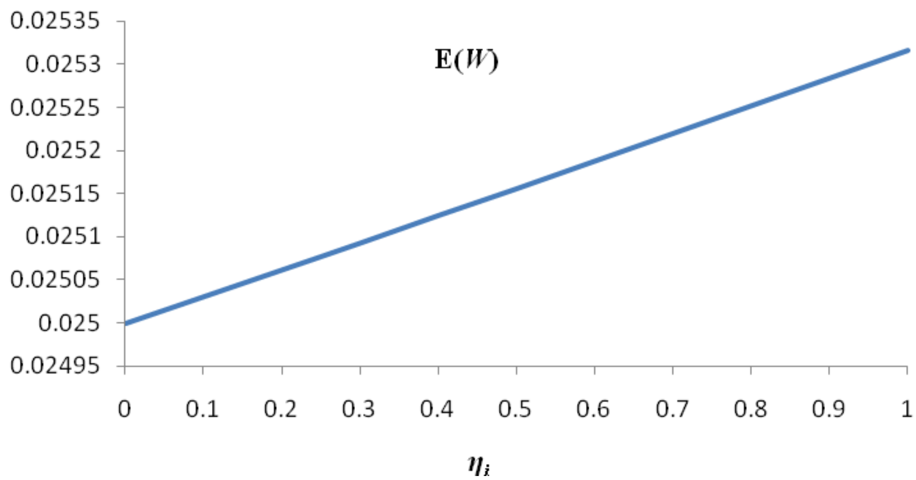

Figure 5.1 The mean of $W(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_i \le 1$.

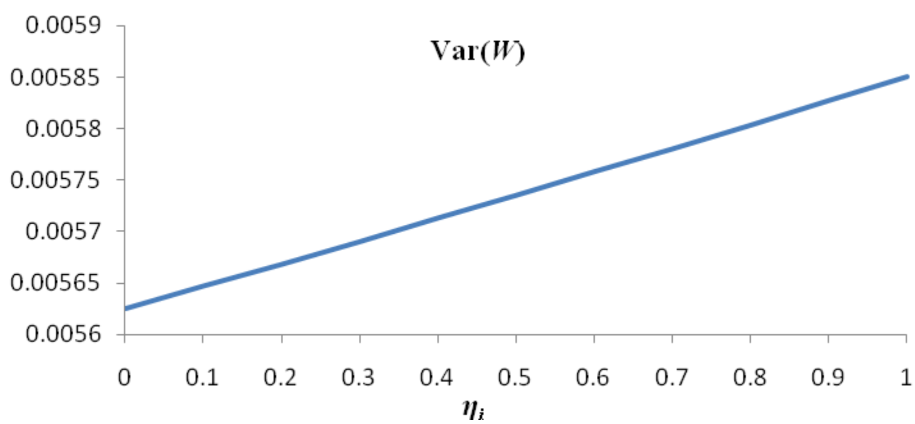

Figure 5.2 The variance of $W(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_i \le 1$.

Figures 5.1 and 5.2 show that both the mean and variance of the waiting time of a normal customer increase in η_i.

Example 3. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 1$ are plotted in Figures 5.3 and 5.4.

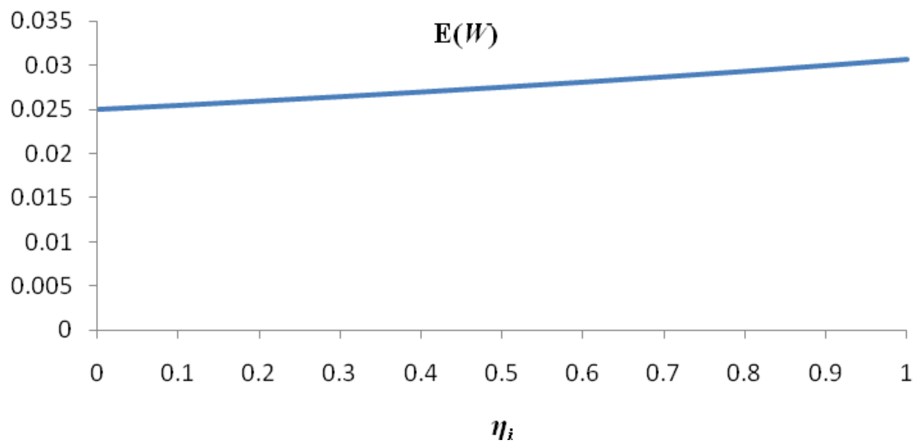

Figure 5.3 The mean of $W(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 1$.

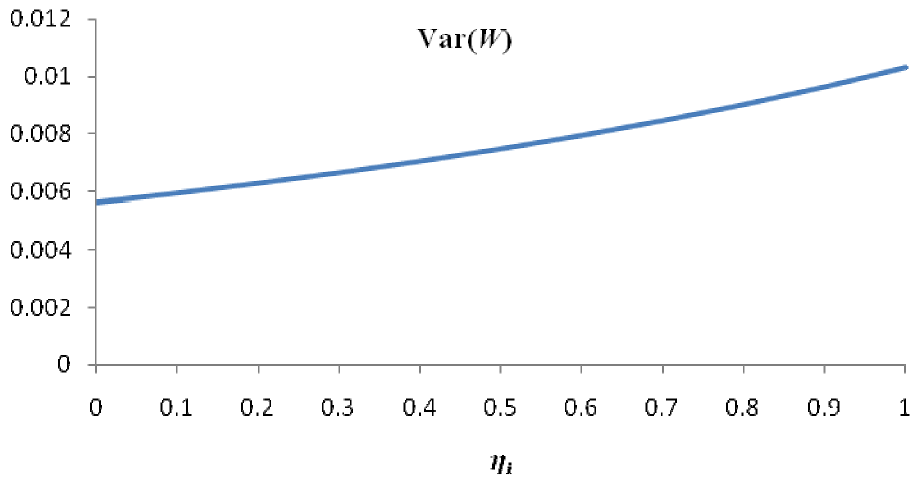

Figure 5.4 The variance of $W(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 1$.

Figures 5.3 and 5.4 show that both the mean and variance of the waiting time of a normal customer increase in η_i.

Example 4. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $W(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$ are plotted in Figures 5.5 and 5.6.

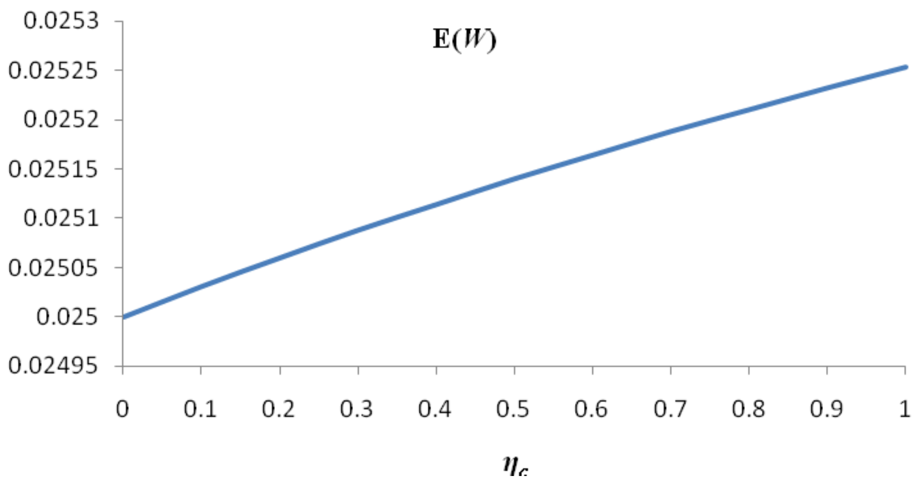

Figure 5.5 The mean of $W(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$.

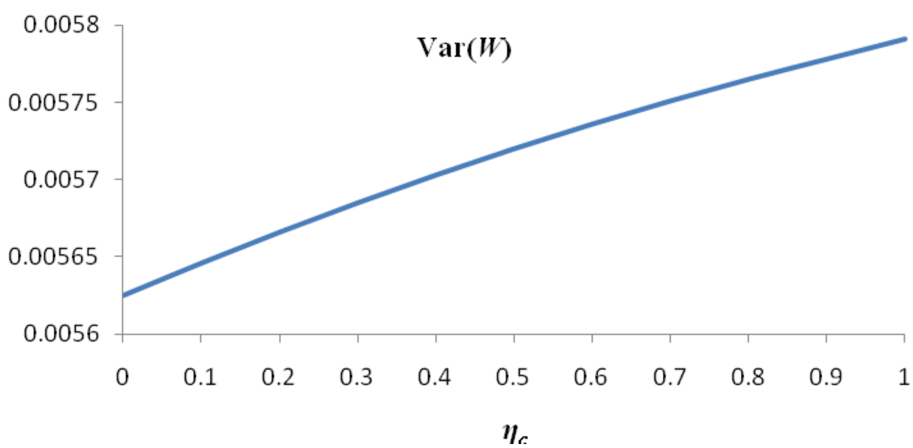

Figure 5.6 The variance of $W(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$.

Figures 5.5 and 5.6 show that both the mean and variance of the waiting time of a normal customer increase in η_c.

Example 5. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $W(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$ are plotted in Figures 5.7 and 5.8.

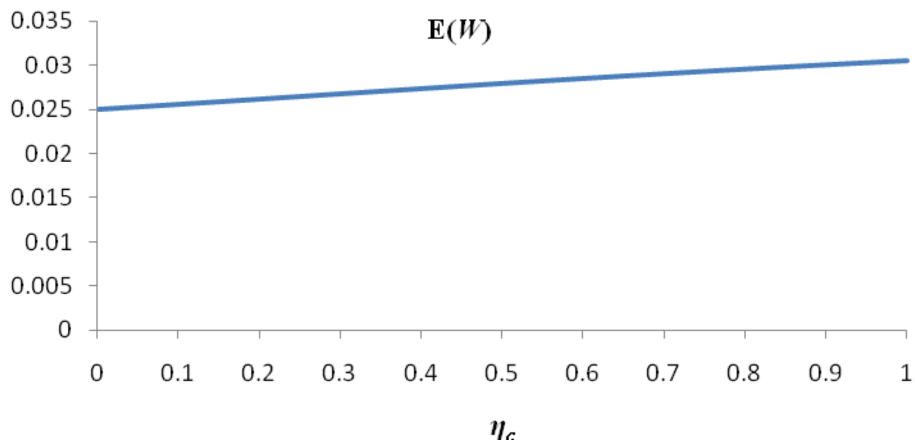

Figure 5.7 The mean of $W(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$.

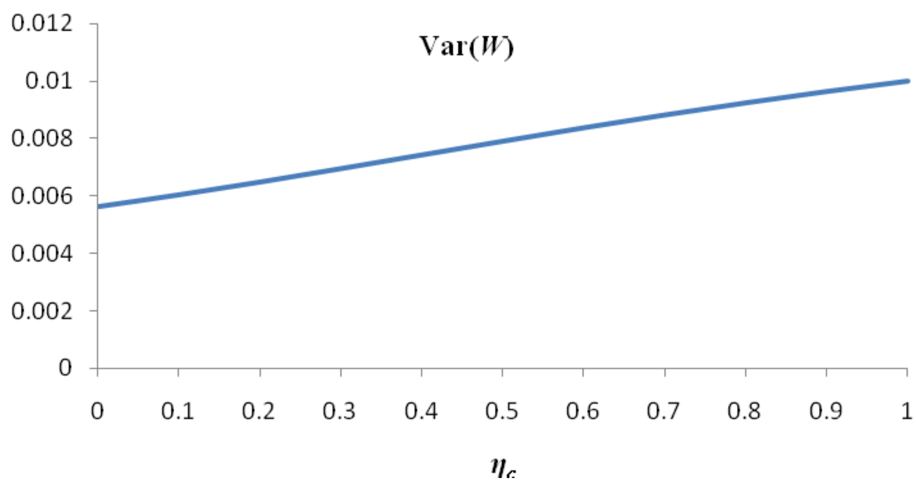

Figure 5.8 The variance of $W(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$.

Figures 5.7 and 5.8 show that both the mean and variance of the waiting time of a normal customer increase in η_c.

Example 6. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $W(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$ are plotted in Figures 5.9 and 5.10.

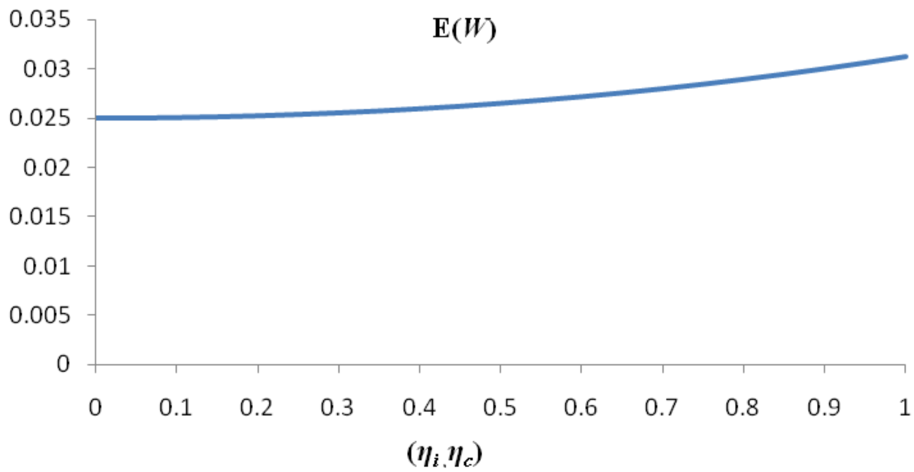

Figure 5.9 The mean of $W(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$.

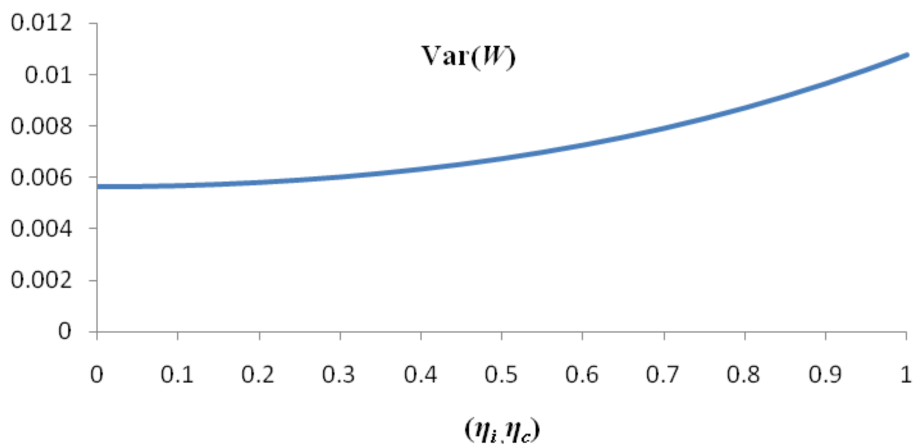

Figure 5.10 The variance of $W(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$.

Figures 5.9 and 5.10 show that both the mean and variance of the waiting time of a normal customer increase in η_i and η_c.

Example 7. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The mean and variance of $W(\eta_i, \eta_c)$ at boundary points are computed in Table 5.2.

Table 5.2 The mean and variance of $W(\eta_i, \eta_c)$ at boundary points when $(\rho = 0.8)$

λ	μ	ρ	η_i	η_c	**E(*W*)**	**Var(*W*)**
8	10	0.8	0	0	0.4	0.24
8	10	0.8	0	1	0.4	0.24
8	10	0.8	1	0	0.4	0.24
8	10	0.8	1	1	2	14

The results in Tables 5.2 for the mean and variance of $W(\eta_i, \eta_c)$ at boundary points correspond with the results that could be obtained from Corollary 4.1.1for boundary points when $\rho = 0.8$. This verifies the accuracy of the computational method in computing the mean and variance of the waiting time of normal customers.

Example 8. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The mean and variance of $W(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_i \le 1$ are plotted in Figures 5.11 and 5.12.

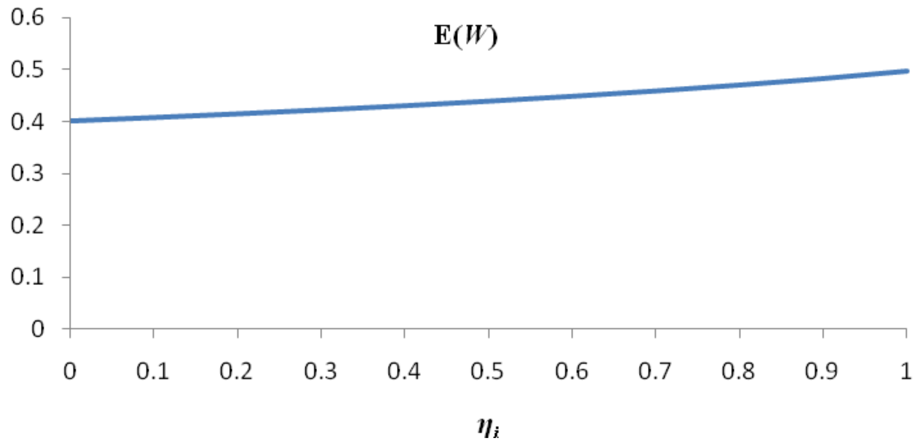

Figure 5.11 The mean of $W(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_i \le 1$.

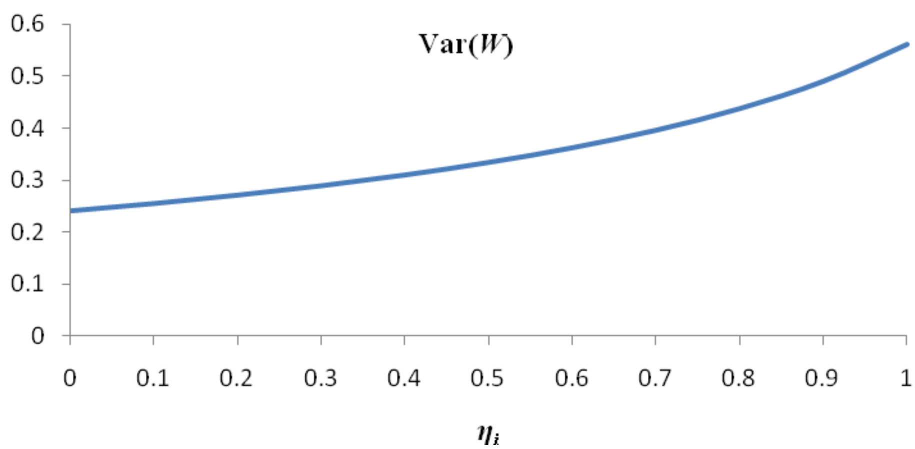

Figure 5.12 The variance of $W(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_i \le 1$.

Figures 5.11 and 5.12 show that both the mean and variance of the waiting time of a normal customer increase in η_i.

Example 9. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The mean and variance of $W(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 1$ are plotted in Figures 5.13 and 5.14.

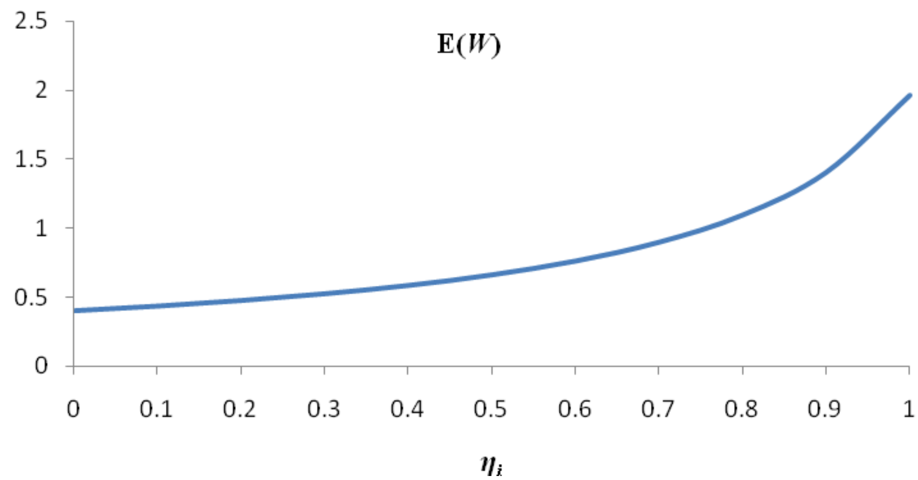

Figure 5.13 The mean of $W(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 1$.

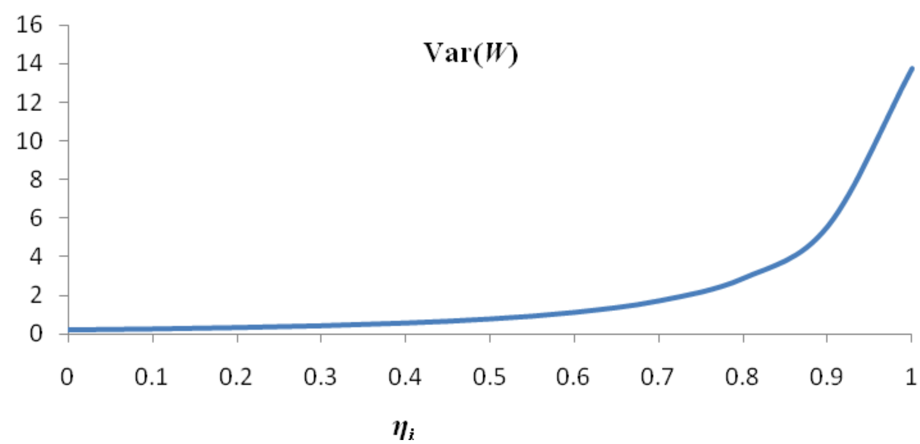

Figure 5.14 The variance of $W(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 1$.

Figures 5.13 and 5.14 show that both the mean and variance of the waiting time of a normal customer increase in η_i.

Example 10. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The mean and variance of $W(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$ are plotted in Figures 5.15 and 5.16.

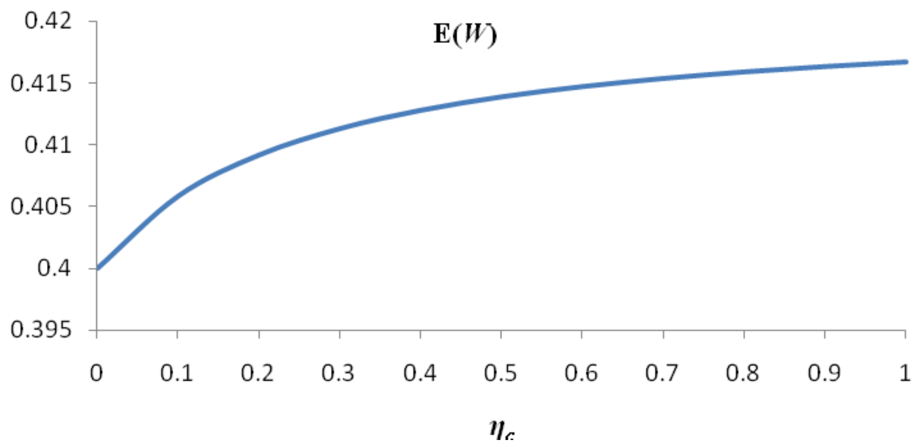

Figure 5.15 The mean of $W(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$.

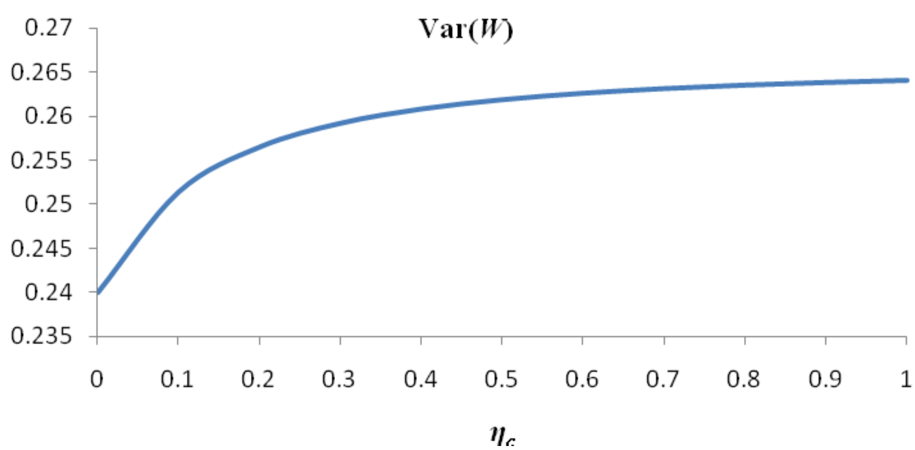

Figure 5.16 The variance of $W(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$.

Figures 5.15 and 5.16 show that both the mean and variance of the waiting time of a normal customer increase in η_c.

Example 11. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The mean and variance of $W(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$ are plotted in Figures 5.17 and 5.18.

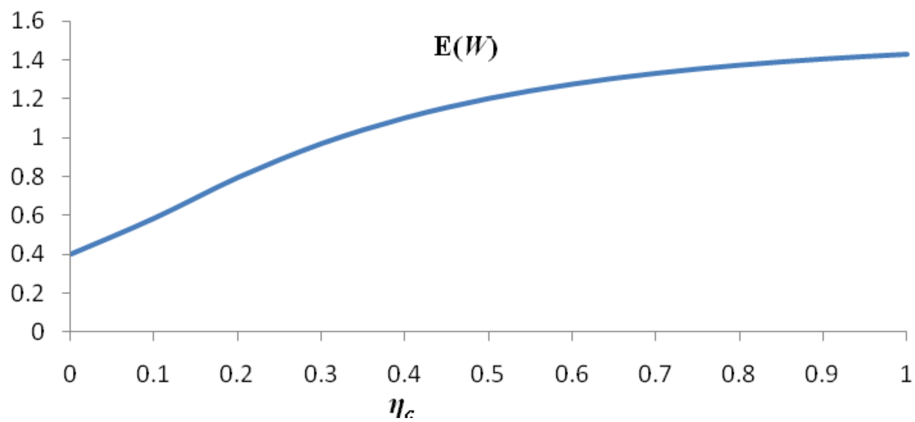

Figure 5.17 The mean of $W(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$.

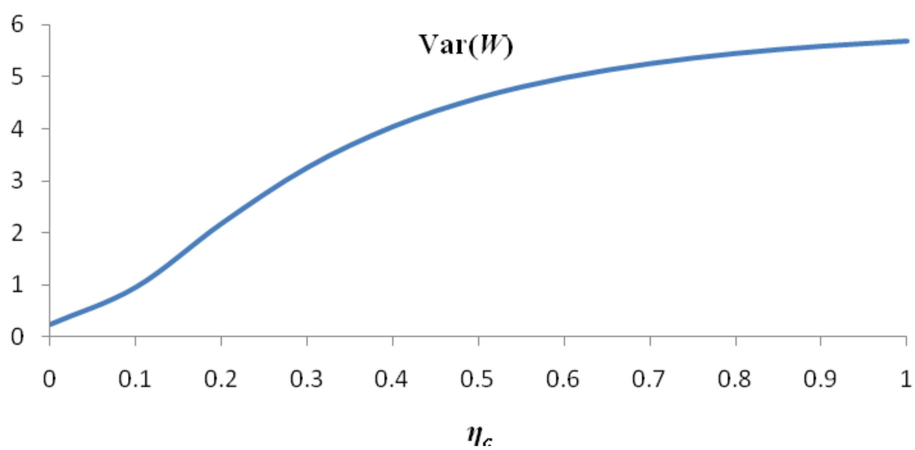

Figure 5.18 The variance of $W(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$.

Figures 5.17 and 5.18 show that both the mean and variance of the waiting time of a normal customer increase in η_c.

Example 12. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The mean and variance of $W(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$ are plotted in Figures 5.19 and 5.20.

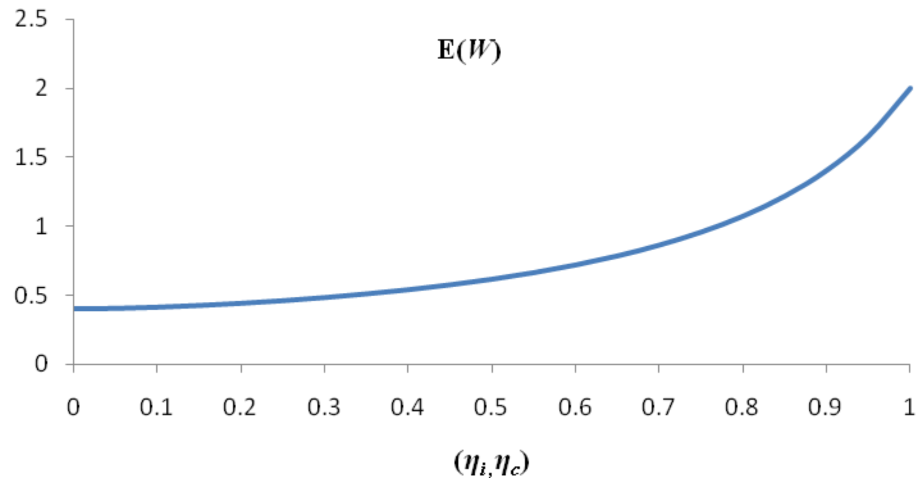

Figure 5.19 The mean of $W(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$.

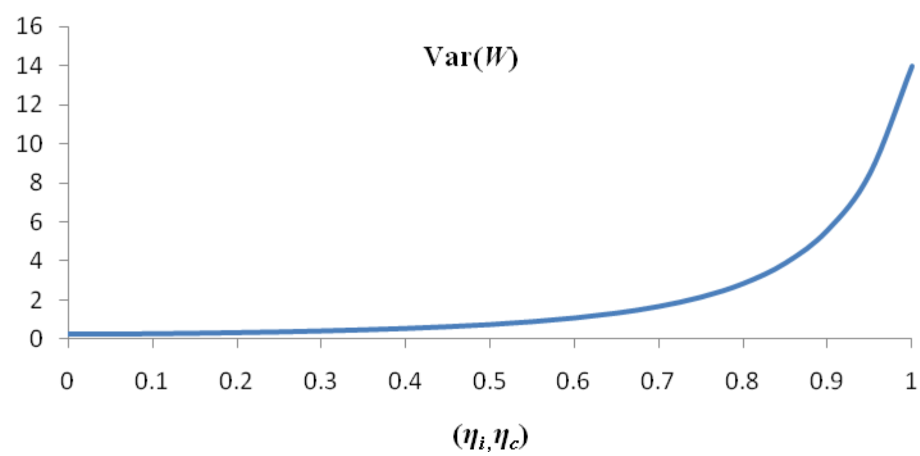

Figure 5.20 The variance of $W(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$.

Figures 5.19 and 5.20 show that both the mean and variance of the waiting time of a normal customer increase in η_i and η_c.

Summary:

The mean and variance of the waiting time of normal customers increase in both η_i and η_c. Thus, the minimums and maximums of the mean and variance are attained at boundary points (η_i, 0), (η_i, 1), $(0, \eta_c)$, $(1, \eta_c)$. Numerical examples presented in Section 5.3 show that the mean and variance of the waiting time of a normal customer can increase drastically due to interjection. They also show that the increase is significant at boundary points. The impact of parameters η_i and η_c on the waiting time of a normal customer can be summarized as follows:

1- If $\eta_i = 0$, no customer interjects. For this case, a small increase in η_i may lead to significant increase in the mean and variance of the waiting time of a normal customer if η_c is large. Thus, if individual tolerance of interjection is high, then a small increase in the intention of interjection can lead to significant increase in the mean waiting time.

2- If $\eta_i = 1$, all customers interject. For this case, a small decrease in η_i can lead to significant decrease in the mean and variance of the waiting time of a normal customer. Thus, if the majority of the customers have the intention to interject, then any decrease in that part can lead to significant decrease in the waiting time of normal customers.

3- If $\eta_c = 0$, no customer tolerates interjection. For this case, a small increase in η_c can lead to a significant increase in the mean and variance of the waiting time of a normal customer. Thus, when the tolerance of interjection is close to zero, then it is worth to eliminate any interjection, since any small tolerance can lead to significant increase in waiting times of normal customers.

4- If $\eta_c = 1$, all customers in queue allow interjection. For this case, a change in η_c lead to change in the mean and variance of the waiting time of a normal customer, which may not be as significant as that at $\eta_c = 0$.

Remark: Both the mean and variance of $W(\eta_i, \eta_c)$ increase in η_i and η_c. Therefore, eliminating customer interjection is always beneficial to normal customers.

5.4 Examples for the Waiting Time of an Arbitrary Interjecting Customer

Example 13. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ at boundary points are computed in Table 5.3.

Table 5.3 The mean and variance of $V(\eta_i, \eta_c)$ at boundary points when $(\rho = 0.2)$

λ	μ	ρ	η_i	η_c	E(V)	Var(V)
2	10	0.2	0	0	0.025	0.005625
2	10	0.2	0	1	0.025	0.005625
2	10	0.2	1	0	0.025	0.005625
2	10	0.2	1	1	0.025	0.007188

The results in Table 5.3 correspond with the results that can be obtained based on equations (4.3.7), (4.3.8), (4.3.9) and (4.3.10) for the mean and variance of the waiting time of an arbitrary interjecting customer at the boundary points when $\rho = 0.2$. This verifies the accuracy of the computational method in computing the mean and variance of the waiting time of interjecting customers.

Example 14. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_i \le 1$ are plotted in Figures 5.21 and 5.22.

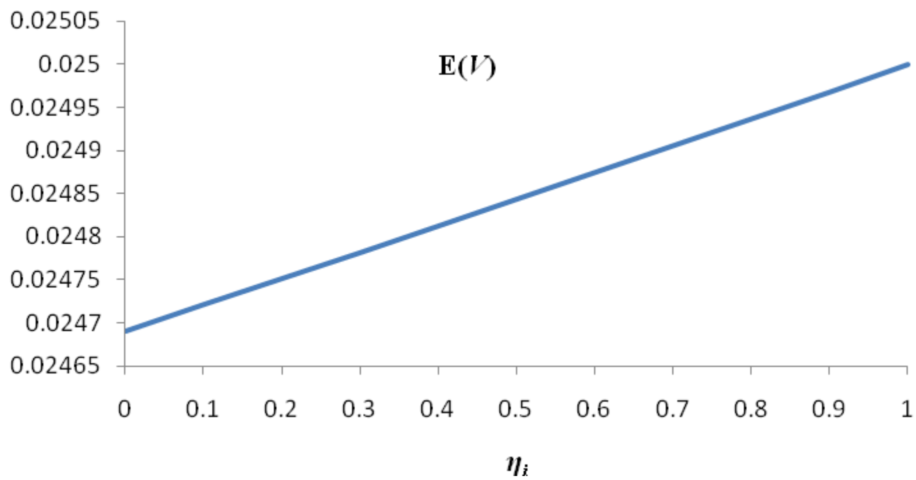

Figure 5.21 The mean of $V(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_i \le 1$.

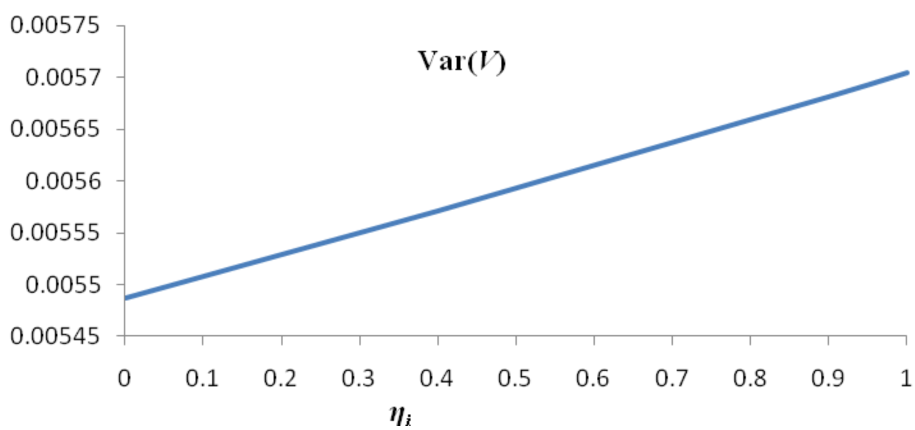

Figure 5.22 The variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_i \le 1$.

Figures 5.21 and 5.22 show that both the mean and variance of the waiting time of an arbitrary interjecting customer increase in η_i.

Example 15. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \leq \eta_i \leq 1$ are plotted in Figures 5.23 and 5.24.

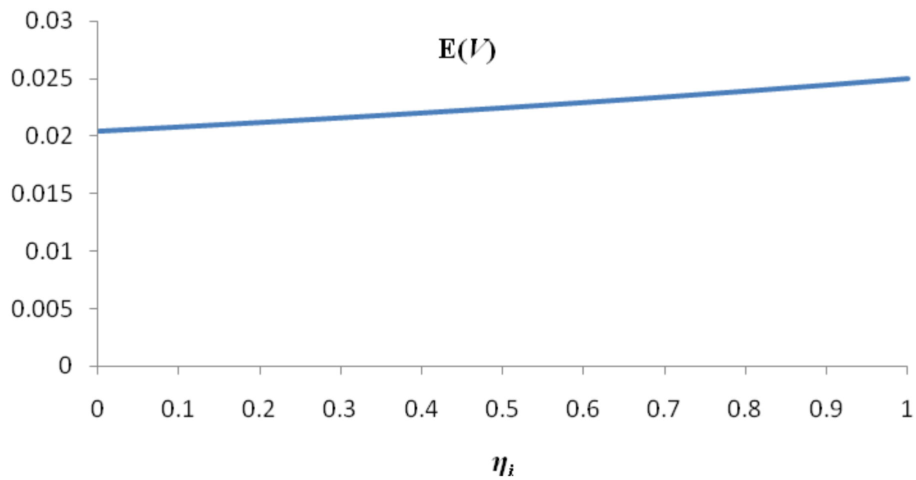

Figure 5.23 The mean of $V(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \leq \eta_i \leq 1$.

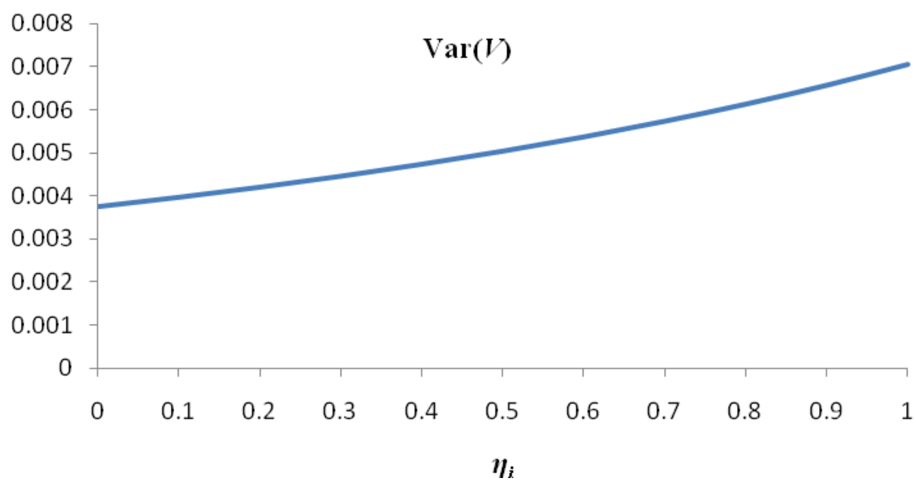

Figure 5.24 The variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \leq \eta_i \leq 1$.

Figures 5.23 and 5.24 show that both the mean and variance of the waiting time of an arbitrary interjecting customer increase in η_i.

Example 16. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$ are plotted in Figures 5.25 and 5.26.

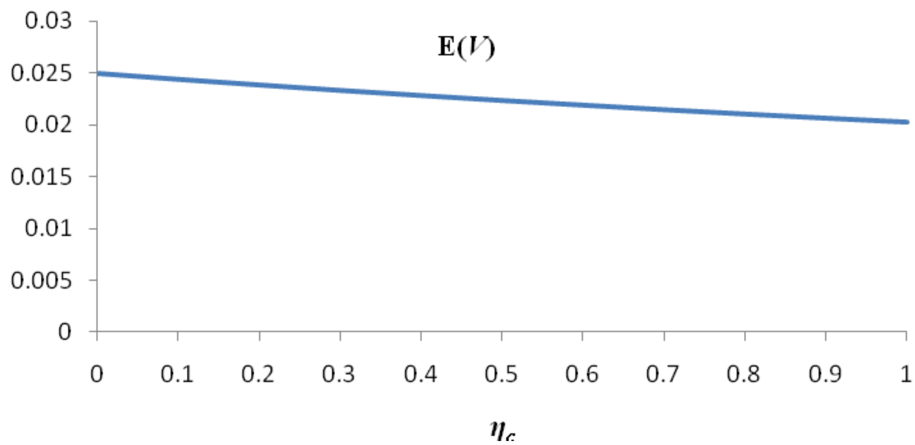

Figure 5.25 The mean of $V(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$.

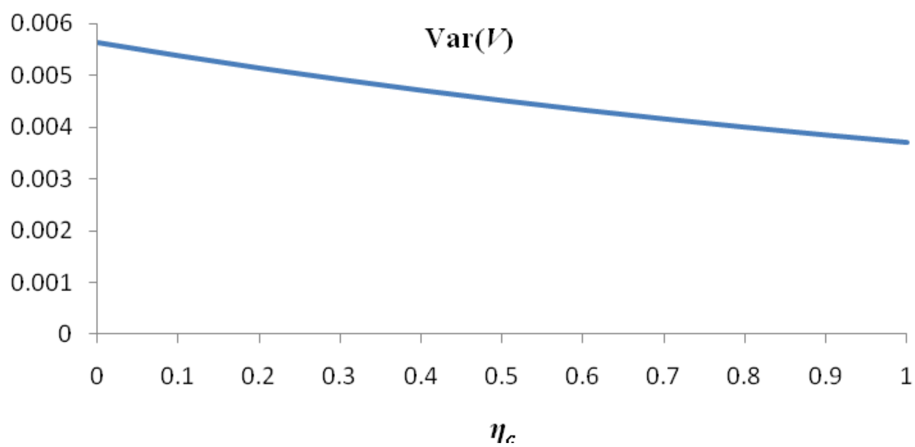

Figure 5.26 The variance of $V(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$.

Figures 5.25 and 5.26 show that both the mean and variance of the waiting time of an arbitrary interjecting customer decrease in η_c when η_i is small and approaches zero.

Example 17. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$ are plotted in Figures 5.27 and 5.28.

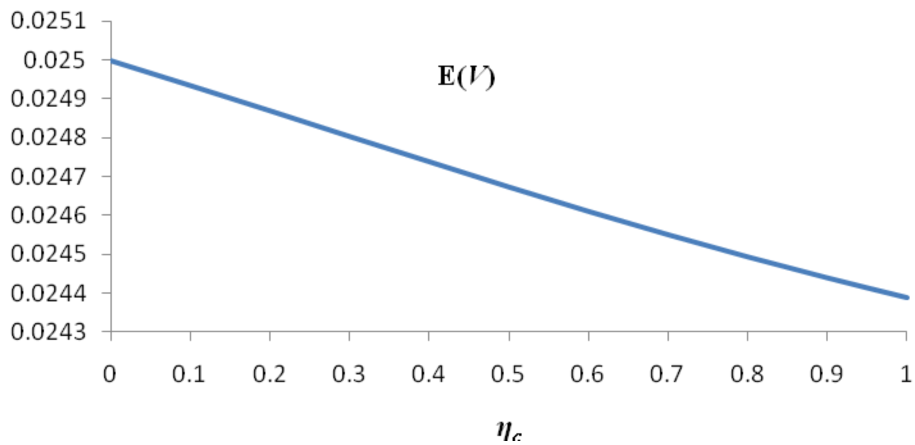

Figure 5.27 The mean of $V(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$.

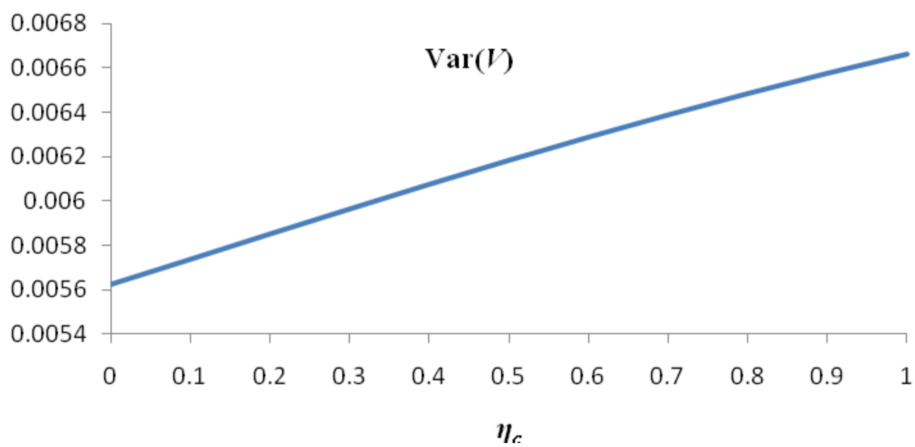

Figure 5.28 The variance of $V(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$.

Figures 5.27 and 5.28 show that the mean of $V(\eta_i, \eta_c)$ decreases in η_c when η_i is large but the variance of $V(\eta_i, \eta_c)$ increases in η_c when η_i is large, which is different from the impact of η_c on the variance of $V(\eta_i, \eta_c)$ when η_i is small and approaches zero.

Example 18. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$ are plotted in Figures 5.29 and 5.30.

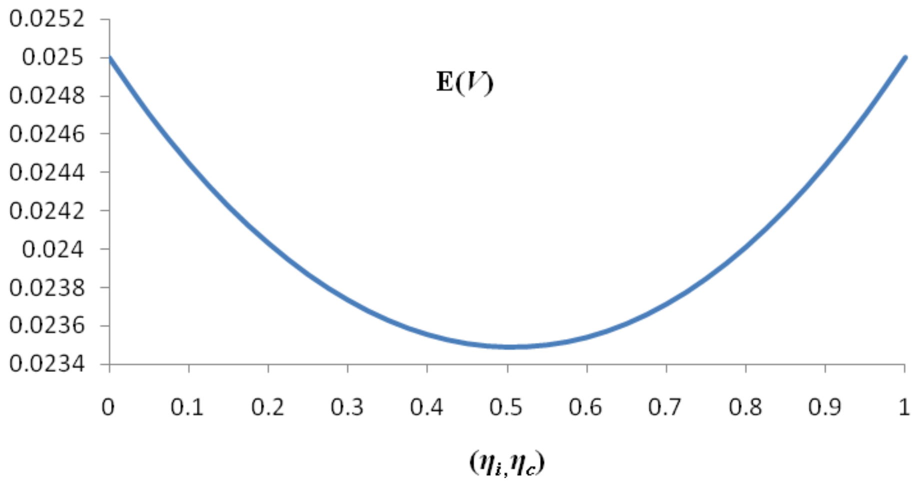

Figure 5.29 The mean of $V(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$.

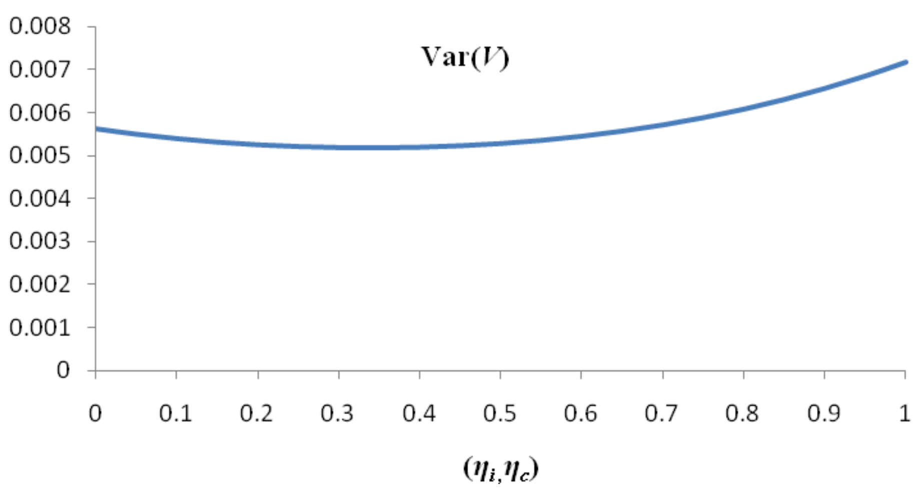

Figure 5.30 The variance of $V(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$.

Figure 5.29 shows that the mean of $V(\eta_i, \eta_c)$ decreases as $\eta_i = \eta_c$ approaches 0.5 and after that, it starts to increase. Figure 5.30 shows that the variance of $V(\eta_i, \eta_c)$ decreases as $\eta_i = \eta_c$ approaches 0.35 and after that, it starts to increase.

Example 19. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ at boundary points are computed in Table 5.4.

Table 5.4 The mean and variance of $V(\eta_i, \eta_c)$ at boundary points when $(\rho = 0.8)$

λ	μ	ρ	η_i	η_c	E(V)	Var(V)
8	10	0.8	0	0	0.4	0.24
8	10	0.8	0	1	0.4	0.24
8	10	0.8	1	0	0.4	0.24
8	10	0.8	1	1	0.4	1.84

The results in Table 5.4 correspond with the results that can be obtained based on equations (4.3.7), (4.3.8), (4.3.9) and (4.3.10) for the mean and variance of the waiting time of an arbitrary interjecting customer at boundary points when $\rho = 0.8$. This verifies the accuracy of the computational method in computing the mean and variance of the waiting time of interjecting customers.

Example 20. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_\mathrm{i} \le 1$ are plotted in Figures 5.31 and 5.32.

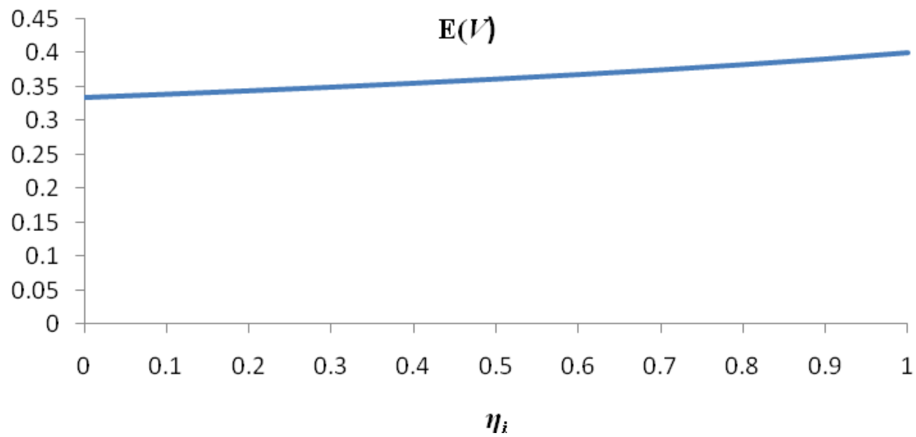

Figure 5.31 The mean of $V(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_\mathrm{i} \le 1$.

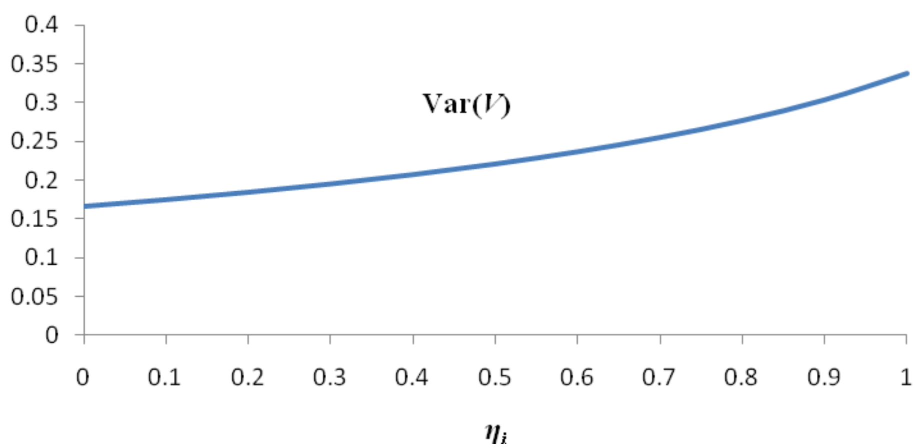

Figure 5.32 The variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_\mathrm{i} \le 1$.

Figures 5.31 and 5.32 show that both the mean and variance of the waiting time of an arbitrary interjecting customer increase in η_i when η_c is small and approaches zero.

Example 21. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 1$ are plotted in Figures 5.33 and 5.34.

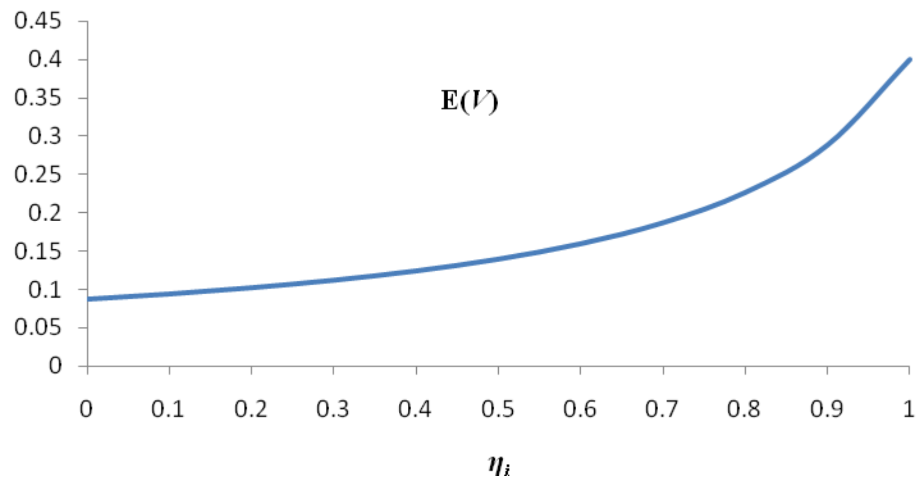

Figure 5.33 The mean of $V(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 1$.

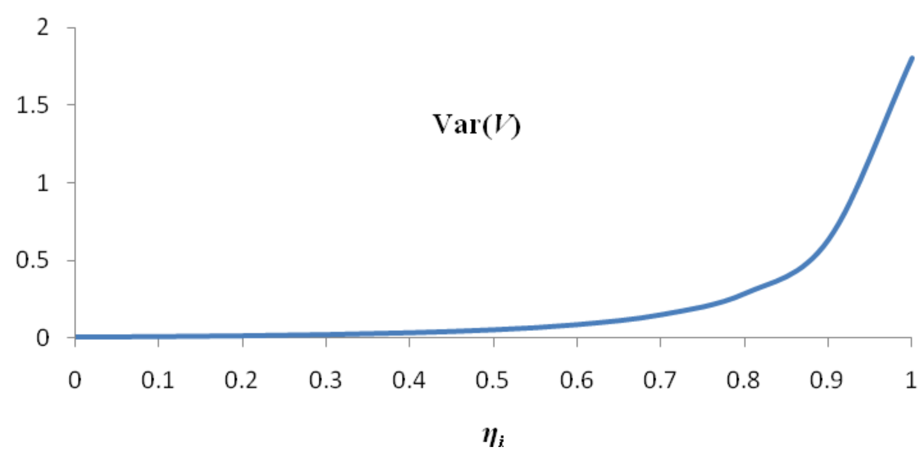

Figure 5.34 The variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 1$.

Figures 5.33 and 5.34 show that both the mean and variance of the waiting time of an arbitrary interjecting customer increase in η_i when η_c is large and approaches one.

Example 22. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$ are plotted in Figures 5.35 and 5.36.

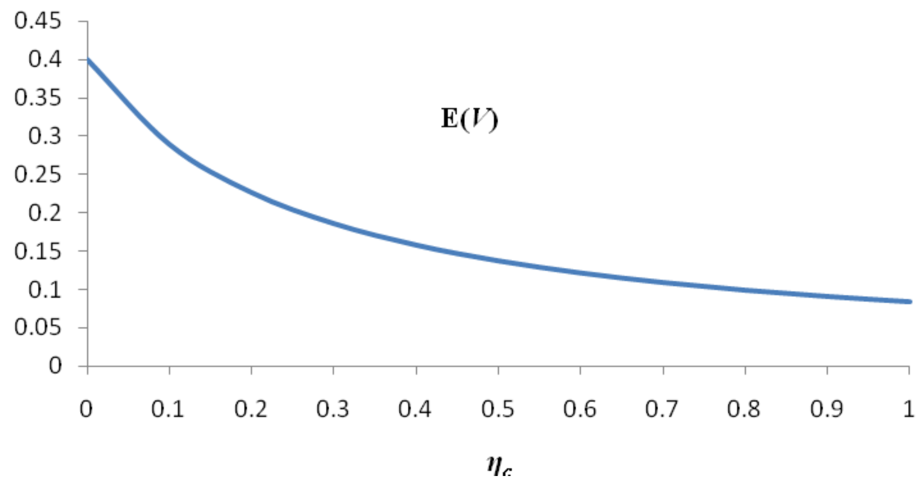

Figure 5.35 The mean of $V(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$.

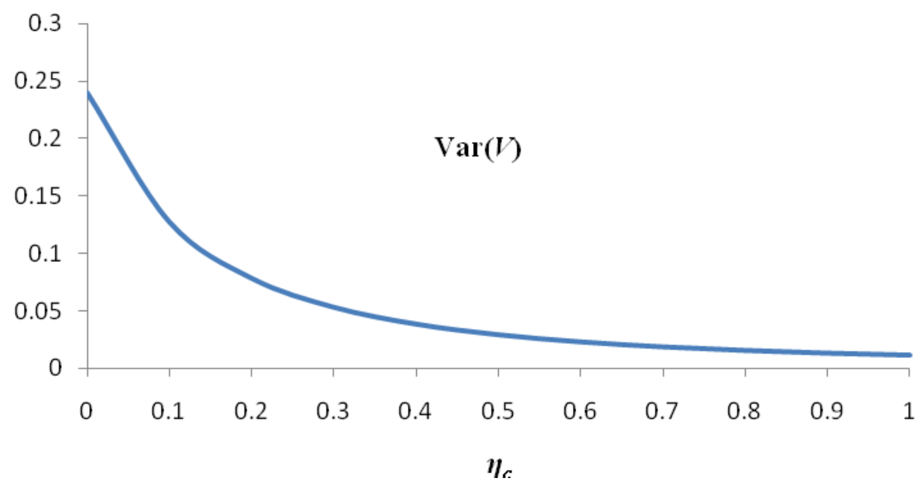

Figure 5.36 The variance of $V(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$.

Figures 5.35 and 5.36 show that both the mean and variance of the waiting time of an arbitrary interjecting customer decrease in η_c when η_i is small and approaches zero.

Example 23. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$ are plotted in Figures 5.37 and 5.38.

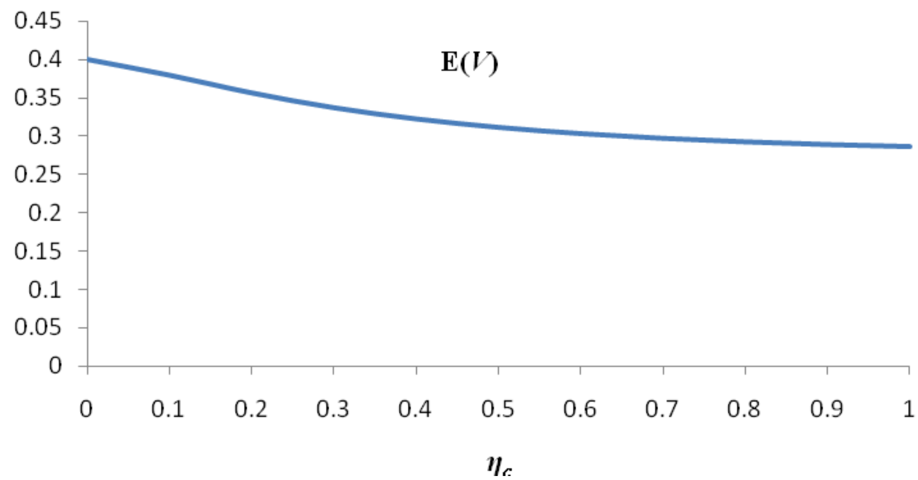

Figure 5.37 The mean of $V(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$.

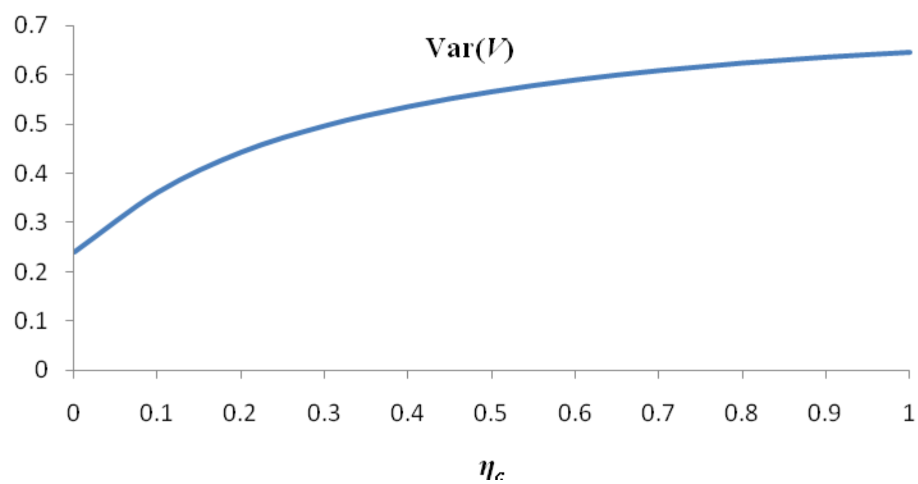

Figure 5.38 The variance of $V(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$.

Figures 5.37 and 5.38 show that the mean of $V(\eta_i, \eta_c)$ decreases in η_c when η_i is large but the variance of $V(\eta_i, \eta_c)$ increases in η_c when η_i is large, which is different from the impact of η_c on the variance of $V(\eta_i, \eta_c)$ when η_i is small and approaches zero.

Example 24. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$ are plotted in Figures 5.39 and 5.40.

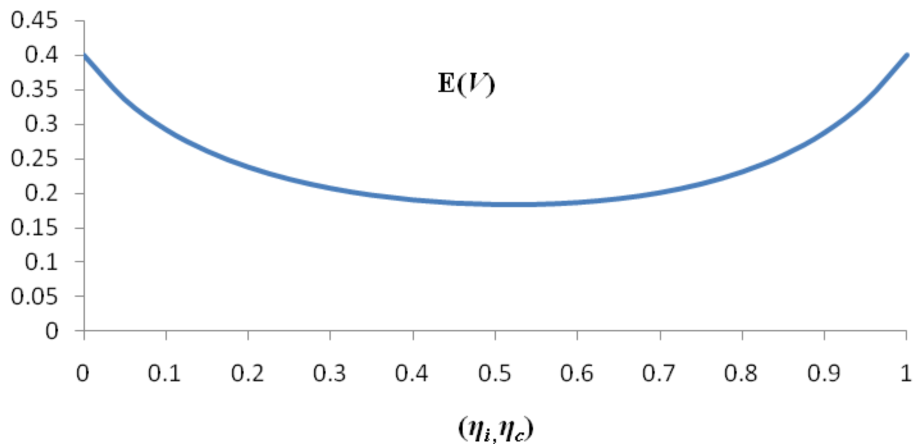

Figure 5.39 The mean of $V(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$.

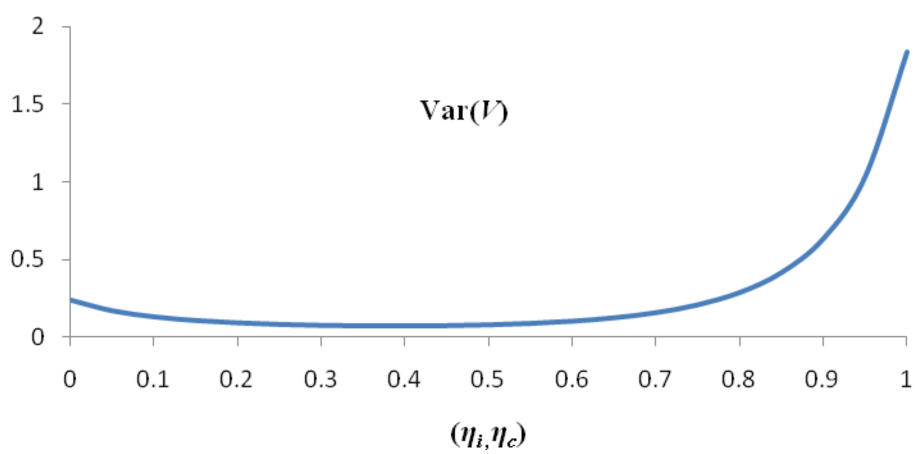

Figure 5.40 The variance of $V(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$.

Figure 5.39 shows that the mean of $V(\eta_i, \eta_c)$ decreases as $\eta_i = \eta_c$ approaches 0.5 and after that, it starts to increase. Figure 5.40 shows that the variance of $V(\eta_i, \eta_c)$ decreases as $\eta_i = \eta_c$ approaches 0.4 and after that, it starts to increase.

Summary:

The mean and variance of the waiting time of an arbitrary interjecting customer increase in η_i and are a monotone function of η_i. It is clear that η_i has the same impact on the waiting time of both normal and interjecting customers. However, reducing η_c will actually increase the waiting time of an interjecting customer, since the possibility for the interjecting customer to cut in at a position close to the server is smaller. This is different from the waiting time of a normal customer. Thus, the mean waiting time of an arbitrary interjecting customer is minimized at $\eta_c = 1$. Also, it can be concluded that the mean of the waiting time is more sensitive at $\eta_c = 0$ than at $\eta_c = 1$. In other words, a small increase in η_c at $\eta_c = 0$ leads to a significant decrease in the mean of the waiting time, but a small decrease in η_c at $\eta_c = 1$ leads to a small increase in the mean of the waiting time. The relationship between the variance of $V(\eta_i, \eta_c)$ and (η_i, η_c) is more complicated. For instance, if η_i is small, the variance decreases in η_c. If η_i is large, the variance increases in η_c. In fact, $Var[V(\eta_i, \eta_c)]$ may not be a monotone function of η_c. Thus, the impact of η_c on the waiting time of an interjecting customer is different from its impact on the waiting time of a normal customer and the impact of η_i on the waiting times. For fixed η_c, the variance increases in η_i, which is consistent with intuition.

Example 25. Consider a queueing model with $\lambda = 12$, $\mu = 10$ and $0 \le \eta_i \le 0.08$. Then $\rho = 1.2 > 1$ and $\eta_i \rho < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_i \le 0.08$ are plotted in Figures 5.41 and 5.42.

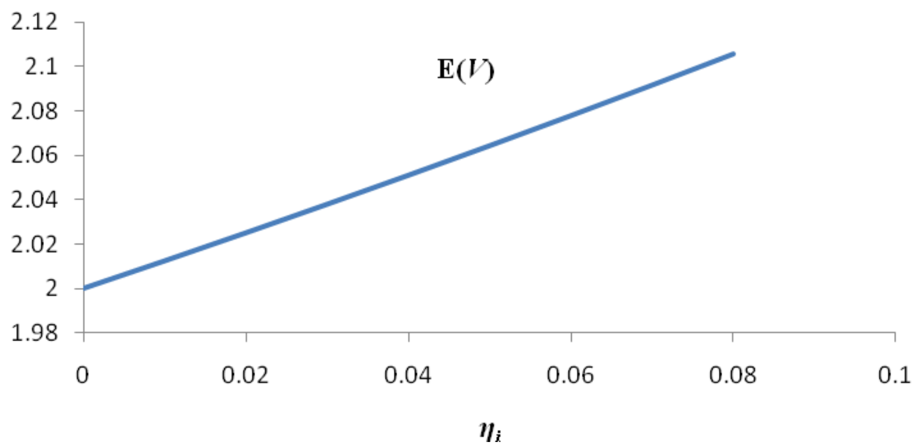

Figure 5.41 The mean of $V(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_\text{i} \le 0.08$.

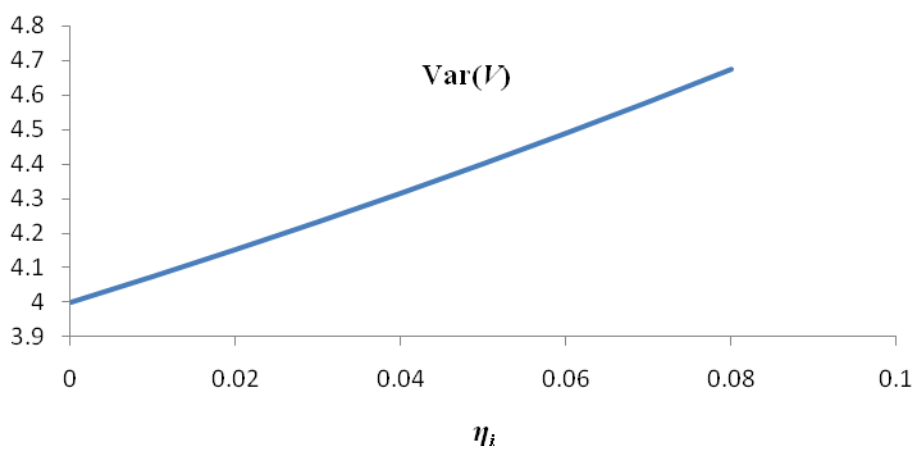

Figure 5.42 The variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_\text{i} \le 0.08$.

Figures 5.41 and 5.42 show that both the mean and variance of $V(\eta_i, \eta_c)$ increases in η_i when η_c is small and approaches zero.

Example 26. Consider a queueing model with $\lambda = 12$, $\mu = 10$ and $0 \le \eta_i \le 0.08$. Then $\rho = 1.2 > 1$ and $\eta_i \rho < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 0.08$ are plotted in Figures 5.43 and 5.44.

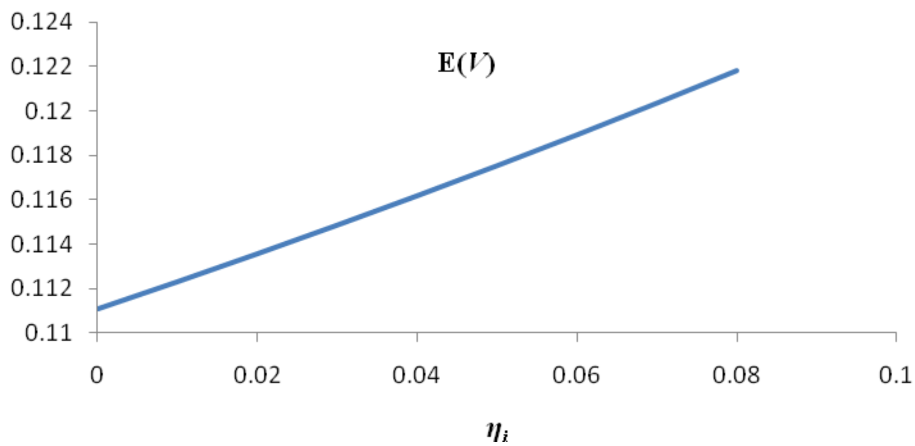

Figure 5.43 The mean of $V(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 0.08$.

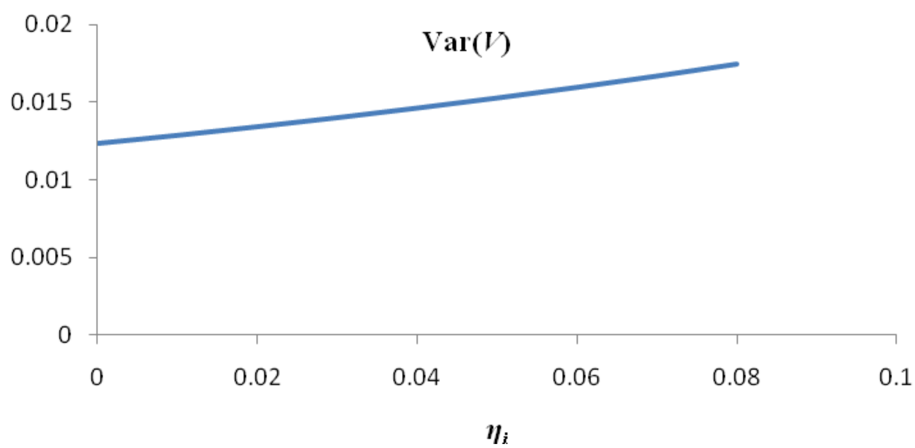

Figure 5.44 The variance of $V(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 0.08$.

Figures 5.43 and 5.44 show that both the mean and variance of $V(\eta_i, \eta_c)$ increase in η_i when η_c is large and approaches one.

Example 27. Consider a queueing model with $\lambda = 12$, $\mu = 10$ and $\eta_i = 0.02$. Then $\rho = 1.2 > 1$ and $\eta_i \rho < 1$. The mean and variance of $V(\eta_i, \eta_c)$ for $\eta_i = 0.02$ and $0 \le \eta_c \le 1$ are plotted in Figures 5.45 and 5.46.

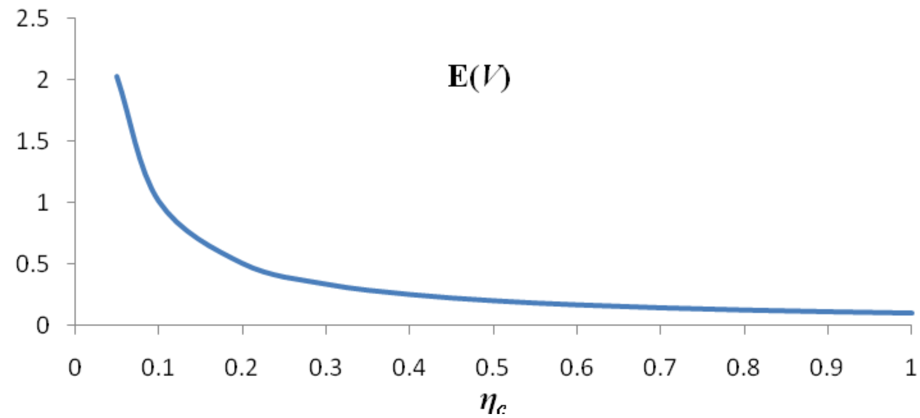

Figure 5.45 The mean of $V(\eta_i, \eta_c)$ for $\eta_i = 0.02$ and $0 \le \eta_c \le 1$.

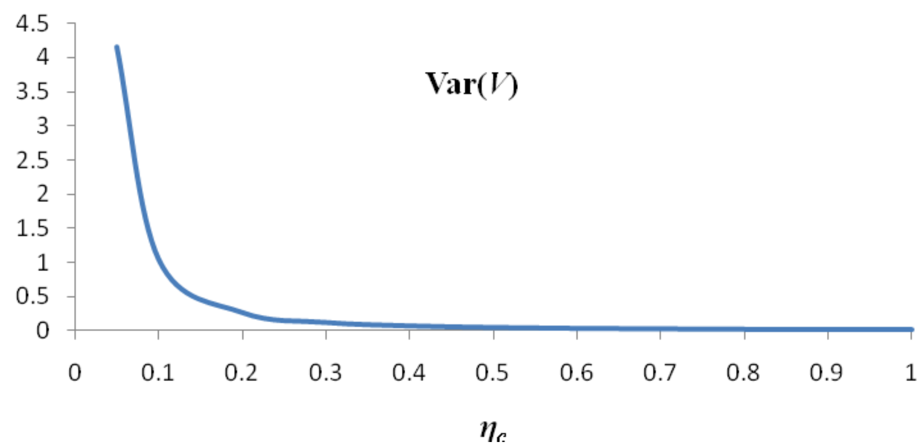

Figure 5.46 The variance of $V(\eta_i, \eta_c)$ for $\eta_i = 0.02$ and $0 \le \eta_c \le 1$.

Figures 5.45 and 5.46 show that the mean and variance of $V(\eta_i, \eta_c)$ decrease in η_c. It is also clear from these figures that the mean and variance of $V(\eta_i, \eta_c)$ approach infinity as

η_c approaches zero. In other words, if $\eta_c = 0$, a small increase in η_c leads to a significant decrease in the mean and variance of $V(\eta_i, \eta_c)$. On the other hand, if $\eta_c = 1$, a small decrease in η_c leads to increase in the mean and variance of $V(\eta_i, \eta_c)$, which is not as significant as that at $\eta_c = 0$.

Summary:

The mean of $V(\eta_i, \eta_c)$ increases in η_i and decreases in η_c when $\rho > 1$ and $\eta_i \rho < 1$. It seems that the mean of $V(\eta_i, \eta_c)$ depends on η_i and η_c in a way similar to the case with $\rho < 1$. However, the variance of $V(\eta_i, \eta_c)$ behaves differently. For this case, the variance decreases in η_c.

5.5 Examples for the Waiting Time of an Arbitrary Customer

Example 28. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The mean and variance of $W_a(\eta_i, \eta_c)$ at boundary points are computed in Table 5.5.

Table 5.5 The mean and variance of $W_a(\eta_i, \eta_c)$ at boundary points when $(\rho = 0.2)$

λ	μ	ρ	η_i	η_c	E(*Wa*)	Var(*Wa*)
2	10	0.2	0	0	0.025	0.005625
2	10	0.2	0	1	0.025	0.005625
2	10	0.2	1	0	0.025	0.005625
2	10	0.2	1	1	0.025	0.007188

Since the results for the mean and variance of $W_a(\eta_i, \eta_c)$ at the boundary points in Table 5.5 correspond with the results that can be computed from equations (4.4.2) and (4.4.3) at the boundary points, then the algorithm for computing the mean and variance of $W_a(\eta_i, \eta_c)$ is accurate.

Example 29. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The variance of $W_a(\eta_i, \eta_c)$ for $\eta_c = 0.05$, $\eta_c = 0.9$ and $0 \le \eta_i \le 1$ is plotted in Figures 5.47 and 5.48.

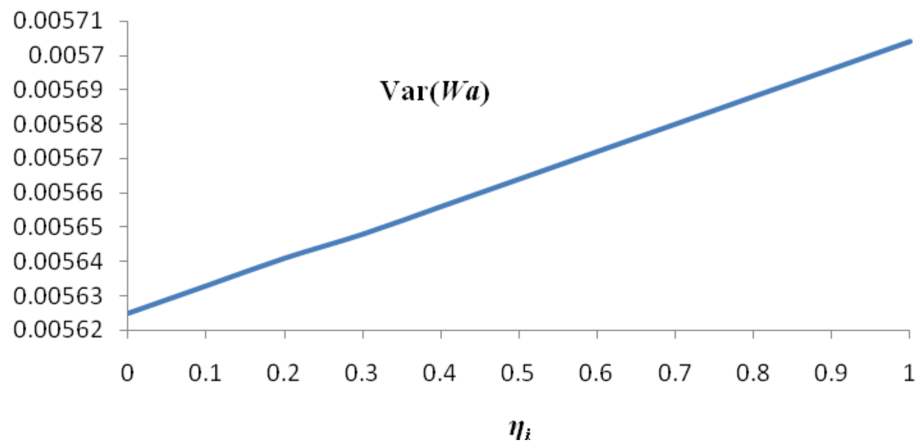

Figure 5.47 The variance of $W_a(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_i \le 1$.

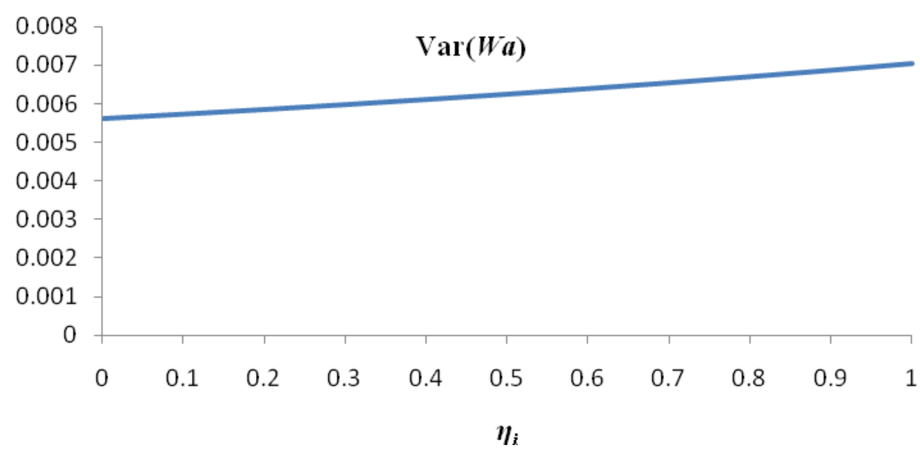

Figure 5.48 The variance of $W_a(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 1$.

Figures 5.47 and 5.48 show that the variance of $W_a(\eta_i, \eta_c)$ increases in η_i for both cases where $\eta_c = 0.05$ and close to zero and, $\eta_c = 0.9$ and close to one.

Example 30. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The variance of $W_a(\eta_i, \eta_c)$ for $\eta_i = 0.05, \eta_i = 0.9$ and $0 \le \eta_c \le 1$ is plotted in Figures 5.49 and 5.50.

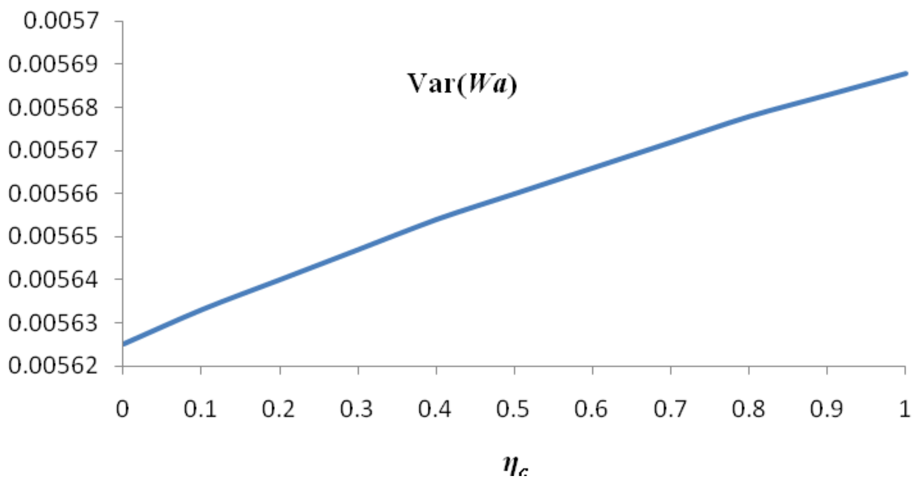

Figure 5.49 The variance of $W_a(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$.

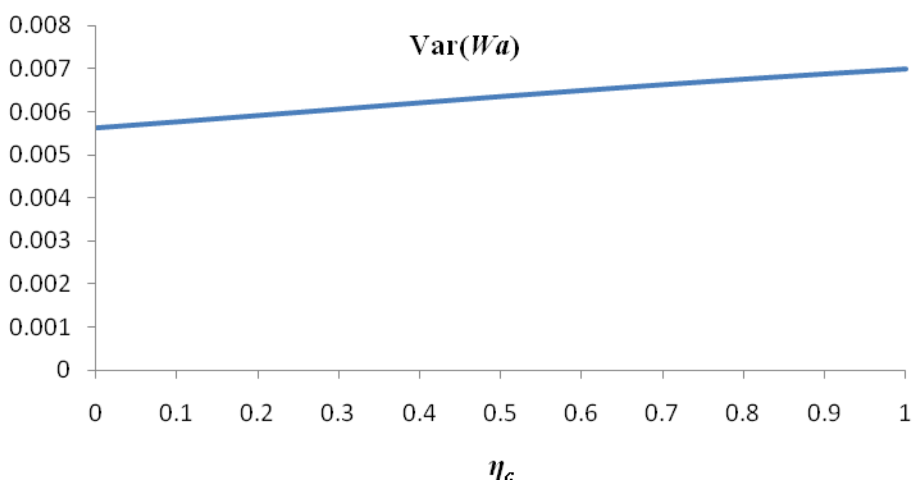

Figure 5.50 The variance of $W_a(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$.

Figures 5.49 and 5.50 show that the variance of $W_a(\eta_i, \eta_c)$ increases in η_c for both cases where $\eta_i = 0.05$ and close to zero and, $\eta_i = 0.9$ and close to one.

Example 31. Consider a queueing model with $\lambda = 2$ and $\mu = 10$. Then $\rho = 0.2 < 1$. The variance of $W_a(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$ is plotted in Figures 5.51.

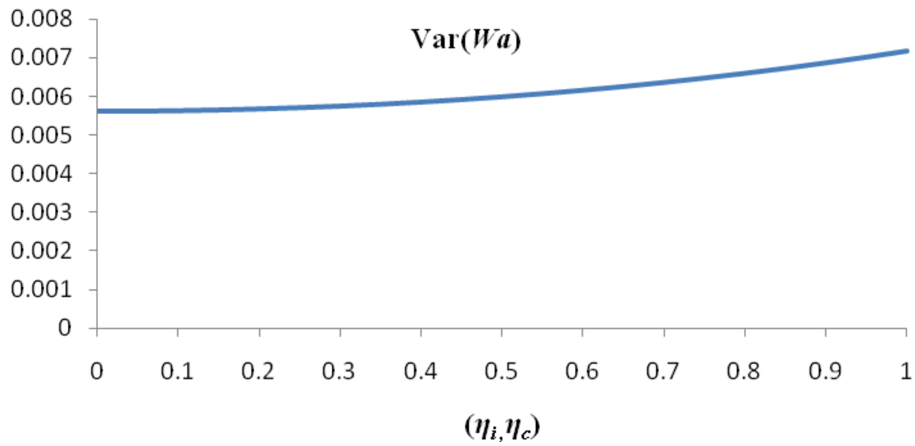

Figure 5.51 The variance of $W_a(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$.

Figure 5.51 shows that the variance of $W_a(\eta_i, \eta_c)$ increases in both η_i and η_c.

Example 32. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The mean and variance of $W_a(\eta_i, \eta_c)$ at boundary points are computed in Table 5.6.

Table 5.6 The mean and variance of $W_a(\eta_i, \eta_c)$ at boundary points when $(\rho = 0.8)$

$\boldsymbol{\lambda}$	$\boldsymbol{\mu}$	$\boldsymbol{\rho}$	$\boldsymbol{\eta_i}$	$\boldsymbol{\eta_c}$	**E(*Wa*)**	**Var(*Wa*)**
8	10	0.8	0	0	0.4	0.24
8	10	0.8	0	1	0.4	0.24
8	10	0.8	1	0	0.4	0.24
8	10	0.8	1	1	0.4	1.84

Since the results for the mean and variance of $W_a(\eta_i, \eta_c)$ at boundary points in Table 5.6 correspond with the results that can be computed from equations (4.4.2) and (4.4.3) at boundary points, then the algorithm for computing the mean and variance of $W_a(\eta_i, \eta_c)$ is accurate.

Example 33. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The variance of $W_a(\eta_i, \eta_c)$ for $\eta_c = 0.05$, $\eta_c = 0.9$ and $0 \le \eta_i \le 1$ is plotted in Figures 5.52 and 5.53.

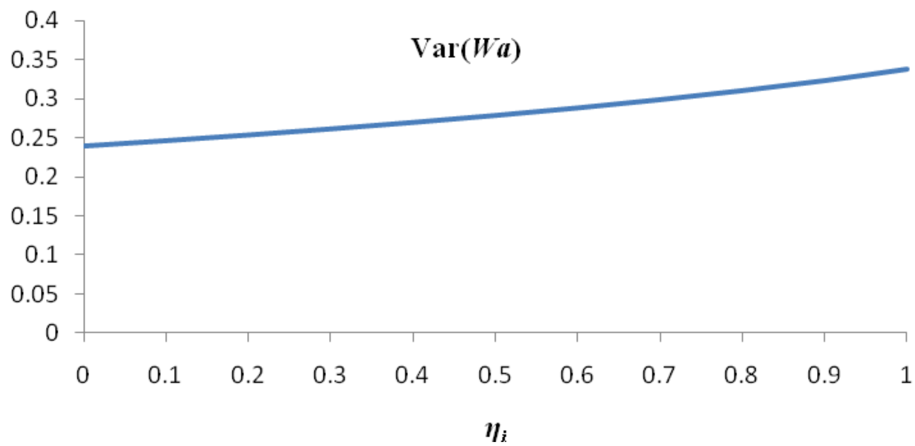

Figure 5.52 The variance of $W_a(\eta_i, \eta_c)$ for $\eta_c = 0.05$ and $0 \le \eta_i \le 1$.

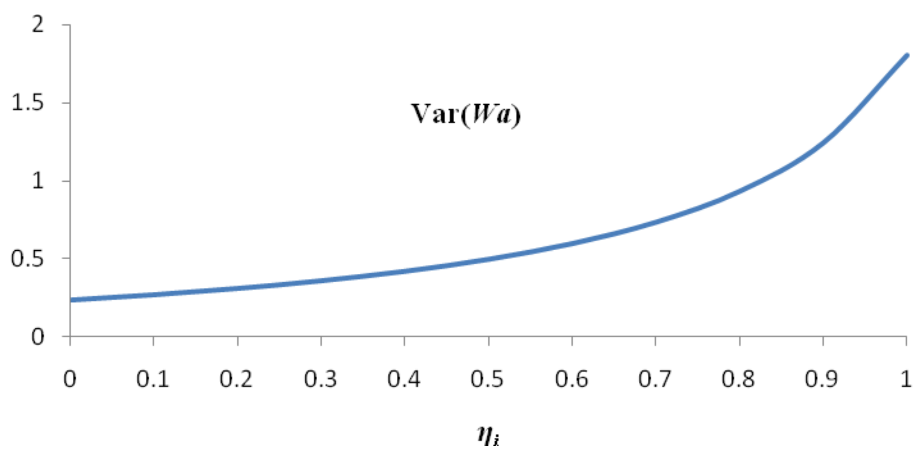

Figure 5.53 The variance of $W_a(\eta_i, \eta_c)$ for $\eta_c = 0.9$ and $0 \le \eta_i \le 1$.

Figures 5.52 and 5.53 show that the variance of $W_a(\eta_i, \eta_c)$ increases in η_i for both cases where $\eta_c = 0.05$ and close to zero and, $\eta_c = 0.9$ and close to one.

Example 34. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The variance of $W_a(\eta_i, \eta_c)$ for $\eta_i = 0.05, \eta_i = 0.9$ and $0 \le \eta_c \le 1$ is plotted in Figures 5.54 and 5.55.

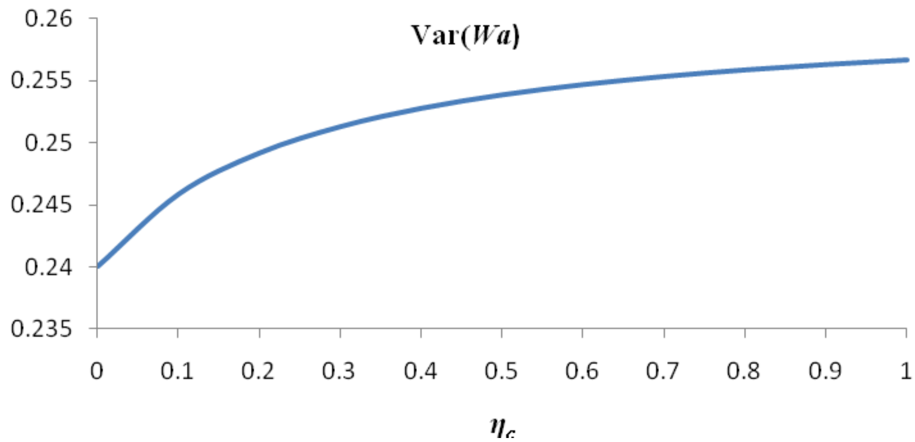

Figure 5.54 The variance of $W_a(\eta_i, \eta_c)$ for $\eta_i = 0.05$ and $0 \le \eta_c \le 1$.

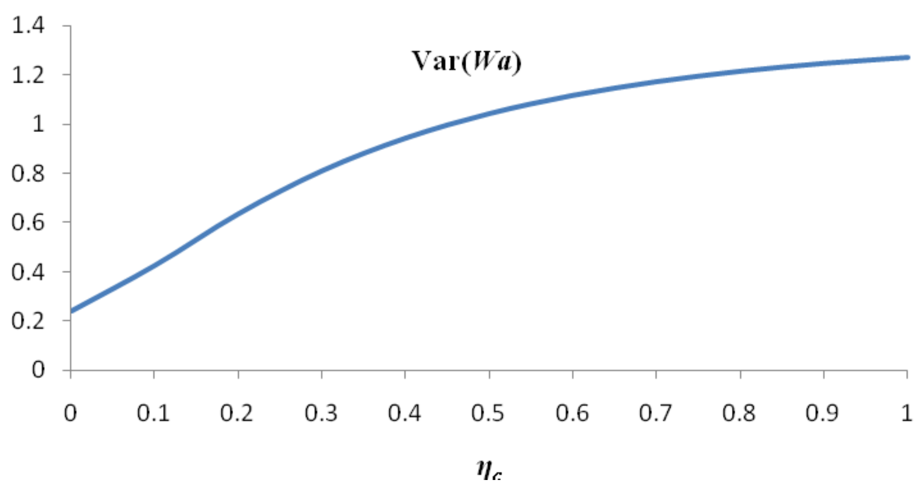

Figure 5.55 The variance of $W_a(\eta_i, \eta_c)$ for $\eta_i = 0.9$ and $0 \le \eta_c \le 1$.

Figures 5.54 and 5.55 show that the variance of $W_a(\eta_i, \eta_c)$ increases in η_c for both cases where $\eta_i = 0.05$ and close to zero and, $\eta_i = 0.9$ and close to one.

Example 35. Consider a queueing model with $\lambda = 8$ and $\mu = 10$. Then $\rho = 0.8 < 1$. The variance of $W_a(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$ is plotted in Figures 5.56.

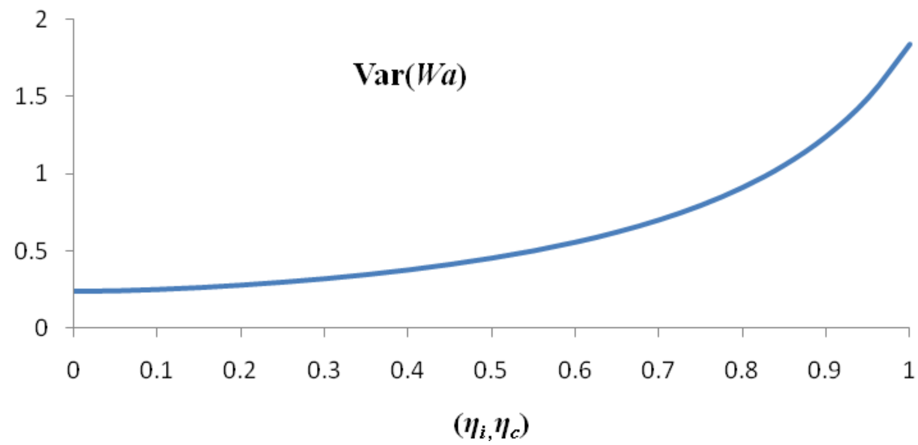

Figure 5.56 The variance of $W_a(\eta_i, \eta_c)$ for $0 \le \eta_i = \eta_c \le 1$.

Figure 5.56 shows that the variance of $W_a(\eta_i, \eta_c)$ increases in both η_i and η_c.

Summary:

As mentioned in Section 4.4, the mean of the waiting time of an arbitrary customer is independent of (η_i, η_c) and is equal to the waiting time of an arbitrary customer in an M/M/1 queue with FCFS service discipline which is equal to $\frac{\rho}{\mu - \lambda}$. However, the variance of $W_a(\eta_i, \eta_c)$ behaves in a way similar to $W(\eta_i, \eta_c)$, which increases in both η_i and η_c, which is different from that of $V(\eta_i, \eta_c)$.

CHAPTER 6

Conclusion

6.1 Summary and Conclusion

In this work, an M/M/1 queueing system with customer interjection has been studied extensively. Customers were classified into two groups: interjecting customers with interjection intention and normal customers who join the queue in the end. Two parameters were introduced to describe the interjection behaviour: the percentage of interjecting customers and the tolerance level of interjection by individual customers who are already waiting in the queue. The impacts of the two parameters on the waiting times of an arbitrary normal customer, an arbitrary interjecting customer, and an arbitrary customer were analyzed in detail. The analysis was focused on the mean and variance of the waiting times, with special attention to boundary points. An efficient algorithm was developed to compute the mean and variance of the waiting times of customers and based on that, numerical examples were developed. Finally, the results of the study were presented.

It was found that the waiting times are sensitive to the tolerance level of interjection by individual customers. For normal customers, an increase in this parameter leads to an increase in the mean and variance of their waiting times. Thus, normal customers would prefer that the tolerance level of interjection by individual customers to be as close to zero as possible.

For interjecting customers, an increase in the tolerance level of interjection by individual customers leads to a decrease in the mean of their waiting times. Depending on the percentage of interjecting customers, the tolerance level of interjection has different impacts on the variance of the waiting time of interjecting customers. If the percentage of interjecting customers is high, then an increase in the tolerance level of interjection by individual customers leads to an increase in the variance of their waiting time in the queue. If the percentage of interjecting customers is small, then an increase in the tolerance level of interjection by individual customers leads to a decrease in the variance of their waiting time in the queue.

For arbitrary customers, it was found that the mean of their waiting times in queue is independent of the percentage of interjecting customers and the tolerance level of interjection by individual customers. However, an increase in either of these two parameters leads to an increase in the variance of the waiting time of arbitrary customers.

It was also found that eliminating customer interjection would be always beneficial to normal customers and arbitrary customers though it would not always be so for interjecting customers.

6.2 Suggestions for Future Work

The research conducted in this work could be expanded in the following directions:

- Study of customer interjection in a more general queueing system like M/G/1, G/M/1 or G/G/1.
- Analysing the same system, but this time by considering the amount of time it takes for an interjecting customer to find a position in queue. Note that this time was negligible in this work.
- Investigating the applications of the model in telecommunications networks.
- In this work, the tolerance level of interjection by individual customers was considered to be constant for all types of customers in the queue. For the future investigations, this parameter could be considered a variable. For instance, a system could be investigated in which different types of people have different levels of tolerance on interjections or, people in different positions have different tolerance levels on customer interjections.

References

[1] I. J. B. F. Adan, J. Wessels, and W. H. M. Zijm. (1991). Analysis of the asymmetric shortest queue problem with threshold jockeying. *Stochastic Models,* volume 7, pp. 615-628.

[2] I. Adiri and I. Domb. (1982, Jan. - Feb.). A single server queueing system working under mixed priority disciplines. *Oper. Res.* 30(1), pp. 97-115. Available: http://www.jstor.org/stable/170313. Retrieved on July 1, 2010.

[3] I. Adiri and I. Domb. (1984, 10). Mixing of non-preemptive and preemptive repeat priority disciplines. *Eur. J. Oper. Res.* 18(1), pp. 86-97.

[4] S. Asmussen. (2003). *Applied probability and queues.* Springer, New York.

[5] B. Avi-Itzhak, I. Brosh, and P. Naor. (1964). On discretionary priority queueing. *Z. Angew., Math. Mech.,* volume 6, pp. 235-242.

[6] U. N. Bhat. (1969, Feb.). Sixty years of queueing theory. *Management Science* 15(6, Application Series), pp. B280-B294. Available: http://www.jstor.org/stable/2628922. Retrieved on July 1, 2010.

[7] H. Chen and D. D. Yao. (2001). *Fundamentals of queueing networks.* Springer, New York.

[8] Y. Z. Cho and C. K. Un. (1993). Analysis of the M/G/1 queue under a combined preemptive/nonpreemptive priority discipline. *Communications, IEEE Transactions on* 41(1), pp. 132-141.

[9] J. W. Cohen. (1982). *The single server queue.* North Holland, Amsterdam.

[10] R. B. Cooper. (1981). *Introduction to queueing theory.* North Holland, Amsterdam.

[11] R. L. Disney and W. E. Mitchell. (1971). A solution for queues with instantaneous jockeying and other customer selection rules. *Naval Res. Logist,* volume 17, pp. 315-325.

[12] S. Drekic and D. A. Stanford. (2000). Threshold-based interventions to optimize performance in preemptive priority queues. *Queueing Systems.* volume 35, pp. 289-315.

[13] S. Drekic and D. A. Stanford. (2001, Jan. - Feb.). Reducing delay in preemptive repeat priority queues. *Oper. Res.* 49(1), pp. 145-156. Available: http://www.jstor.org/stable/222961. Retrieved on July 1, 2010.

[14] E. A. Elsayed and A. Bastani. (1985). General solutions of the jockeying problem. *Eur. J. Oper. Res.* 22(3), pp. 387-396.

[15] E. S. Gordon. (1987). New problems in queues: Social injustice and server production management. Ph. D. Thesis, Massachusetts Institute of Technology.

[16] D. Gross and C. M. Harris. (1985). *Fundamentals of queueing theory.* John Wiley & Sons, New York.

[17] F. A. Haight. (1958). Two queues in parallel. *Biometrika,* volume 45, pp. 401-410.

[18] Qi-Ming. He and M. F. Neuts. (2002). Two M/M/1 queues with transfers of customers. *Queueing Systems,* volume 42, pp. 377-400.

[19] N. K. Jaiswal. (1968). *Priority queues.* Academic Press, New York.

[20] E. P. C. Kao and C. Lin. (1990). A matrix-geometric solution of the jockeying problem. *Eur. J. Oper. Res.* 44(1), pp. 67-74.

[21] D. G. Kendall. (1953, Sep.). Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain. *The Annals of Mathematical Statistics* 24(3), pp. 338-354. Available: http://www.jstor.org/stable/2236285. Retrieved on July 1, 2010.

[22] L. Kleinrock. (1964). A delay dependent queue discipline. *Naval Res. Logistics Quart,* volume 11, pp. 329-341.

[23] L. Kleinrock. (1975). *Queueing systems, volume I: Theory.* John Wiley & Sons, New York.

[24] L. Kleinrock. (1976). *Queueing systems, volume II: Computer applications.* John Wiley & Sons, New York.

[25] R. C. Larson. (1987, Nov. - Dec.). Perspectives on queues: Social justice and the psychology of queueing. *Oper. Res.* 35(6), pp. 895-905. Available: http://www.jstor.org/stable/171439. Retrieved on July 1, 2010.

[26] A. W. Marshall and I. Olkin. (1967). *Inequalities: Theory of majorization and its applications.* Academic Press, New York.

[27] M. F. Neuts. (1981). *Matrix-Geometric solutions in stochastic models: An algorithmic approach.* The Johns Hopkins University Press, Baltimore.

[28] M. F. Neuts. (1986, Dec.). Generalizations of the Pollaczek-Khinchin integral equation in the theory of queues. *Adv. Appl. Probability* 18(4), pp. 952-990. Available: http://www.jstor.org/stable/1427258. Retrieved on July 1, 2010.

[29] M. Paterok and M. Ettl. (1994, Nov. - Dec.). Sojourn time and waiting time distributions for M/GI/1 queues with preemption-distance priorities. *Oper. Res.* 42(6), pp. 1146-1161. Available: http://www.jstor.org/stable/171992. Retrieved on July 1, 2010.

[30] S. Ross. (1983). *Stochastic processes.* John Wiley and Sons, New York.

[31] T. L. Satty. (1966). Seven more years of queues: A lament and a bibliography. *Naval Res. Logistics Quart,* volume 13, pp. 447-476.

[32] E. Seneta. (2006). *Non-negative matrices and Markov chains*. Springer, USA, 2006.

[33] M. Shaked and J. G. Shanthikumar. (2006). *Stochastic orders.* Springer, New York.

[34] H. Takagi. (1990). *Queueing analysis: A foundation of performance evaluation, volume 1: Vacation and priority systems, Part 1.* Elsevier Science Publisher, B. V., Amsterdam.

[35] H. Takagi and Y. Kodera. (1996). Analysis of preemptive loss priority queues with preemption distance. *Queueing Systems,* volume 22, pp. 367-381.

[36] T. Takine and T. Hasegawa. (1994). The workload in the MAP/G/1 queue with state-dependent services its application to a queue with pre-emptive resume priority. *Stochastic Models,* volume 10, pp. 183-204.

[37] J. A. White, J. W. Schmidt, and G. K. Bennett. (1975). *Analysis of queueing systems.* Academic Press, New York.

[38] W. Whitt. (1984). The amount of overtaking in a network of queues. *Networks,* volume 14, pp. 411-426.

[39] W. Whitt. (1986, Jan. - Feb.). Deciding which queue to join: Some counterexamples. *Oper. Res.* 34(1), pp. 55-62. Available: http://www.jstor.org/stable/170670. Retrieved on July 1, 2010.

[40] S. Zeltyn, Z. Feldman, and S. Wasserkrug. (2009). Waiting and sojourn times in a multi-server queue with mixed priorities. *Queueing Systems,* volume 61, pp. 305-328.

[41] Y. Zhao. (1990). Shortest queue models. Ph. D. Thesis, University of Saskatchewan.

[42] Y. Zhao and W. K. Grassmann. (1990). A solution of the shortest queue model with jockeying-in terms of traffic intensity ρ. *Naval Res. Logist,* volume 37, pp. 773-787.

[43] Y. Zhao and W. K. Grassmann. (1995, May - Jun.). Queueing analysis of a jockeying model. *Oper. Res.* 43(3), pp. 520-529. Available: http://www.jstor.org/stable/171875. Retrieved on July 1, 2010.

Appendix A

Stochastic Ordering

Because of the nature of random variables, it is necessary to have a method to compare different random variables with each other. Stochastic ordering is the method to compare random variables. Different researchers have done extensive work about stochastic ordering. This appendix is a brief summary of discussions in Ross [30]. The purpose of this appendix is to provide the reader with the concept of stochastic ordering.

The random variable X is said to be stochastically larger than the random variable Y, written $X \geq_{st} Y$, if

$$P\{X > a\} \geq P\{Y > a\} \quad \text{for all } a. \tag{A.1}$$

Lemma A.1

If $X \geq_{st} Y$, then $E[X] \geq_{st} E[Y]$.

The reader is referred to Ross [30], Chapter 9, for the proof of this lemma.

Proposition A.1

$$X \geq_{st} Y \Leftrightarrow E[f(X)] \geq_{st} E[f(Y)] \text{ for all increasing functions } f.$$

The reader is referred to Ross [30], Chapter 9, for the proof of this proposition.

Appendix B

Proofs

B.1. The proof of Corollary 4.1.4

The moments of $W_n(\eta_i,\eta_c)$ can be obtained from its Laplace transform. Remember from equation (4.1.6) that

$$w_n^*(s) = 1 - \frac{s}{\mu}\left[\sum_{m=1}^{n}\sum_{k=m}^{\infty}\left[(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\right]w_k^*(s)\right].$$

For the first moment of $W_n(\eta_i,\eta_c)$, the following relationship exists:

$$E[W_n(\eta_i,\eta_c)] = -\frac{dw_n^*(s)}{ds}\Big|_{s=0}. \tag{B.1}$$

From equation (4.1.6), it is easy to find that

$$-\frac{dw_n^*(s)}{ds} = +\frac{1}{\mu}\left[\sum_{m=1}^{n}\sum_{k=m}^{\infty}\left[(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\right]w_k^*(s)\right]$$

$$+\frac{s}{\mu}\left[\sum_{m=1}^{n}\sum_{k=m}^{\infty}\left[(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\right]\frac{dw_k^*(s)}{ds}\right]. \tag{B.2}$$

From equations (B.1) and (B.2), the following relationship can be obtained:

$$E[W_n(\eta_i,\eta_c)] = -\frac{dw_n^*(s)}{ds}\Big|_{s=0}$$

$$= +\frac{1}{\mu}\left[\sum_{m=1}^{n}\sum_{k=m}^{\infty}\left[(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\right]w_k^*(0)\right]. \tag{B.3}$$

Since $w_k^*(0)=1$ for all k, then $E[W_n(\eta_i,\eta_c)]$ can be obtained from equation (B.3) as follows:

$$E[W_n(\eta_i,\eta_c)] = \frac{1}{\mu}\sum_{m=1}^{n}\sum_{k=m}^{\infty}(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j}). \tag{B.4}$$

For the second moment of $W_n(\eta_i,\eta_c)$, the following relationship exists:

$$E[W_n^2(\eta_i,\eta_c)] = \frac{d^2 w_n^*(s)}{ds^2}\Big|_{s=0}. \tag{B.5}$$

From equation (4.1.6), it is easy to find that

$$\begin{aligned}\frac{d^2 w_n^*(s)}{ds^2} &= \frac{1}{\mu}\left[\sum_{m=1}^{n}\sum_{k=m}^{\infty}\left[(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\right]\left(-\frac{dw_k^*(s)}{ds}\right)\right] \\ &+ \frac{1}{\mu}\left[\sum_{m=1}^{n}\sum_{k=m}^{\infty}\left[(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\right]\left(-\frac{dw_k^*(s)}{ds}\right)\right] \\ &- \frac{s}{\mu}\left[\sum_{m=1}^{n}\sum_{k=m}^{\infty}\left[(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\right]\frac{d^2 w_k^*(s)}{ds^2}\right].\end{aligned} \tag{B.6}$$

Since $(-\frac{dw_k^*(s)}{ds}\big|_{s=0}) = E[W_k(\eta_i,\eta_c)]$, then the second moment of $W_n(\eta_i,\eta_c)$ can be obtained from equation (B.6) as follows:

$$\begin{aligned}E[W_n^2(\eta_i,\eta_c)] &= \frac{d^2 w_n^*(s)}{ds^2}\Big|_{s=0} \\ &= \frac{2}{\mu}\sum_{m=1}^{n}\sum_{k=m}^{\infty}(\rho\eta_i)^{k-m}\Big(\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\Big)E[W_k(\eta_i,\eta_c)].\end{aligned} \tag{B.7}$$

This completes the proof of Corollary 4.1.4.

B.2. The proof of equation (4.2.2)

It is easy to find the second moment of $W(\eta_i,\eta_c)$ from equation (4.2.1). However, the first moment of $W(\eta_i,\eta_c)$ still needs more discussion.

Remember from equation (4.2.1) that

$$w^*(s) = E[\exp\{-sW(\eta_i,\eta_c)\}] = \sum_{n=0}^{\infty}(1-\rho)\rho^n w_n^*(s), s \geq 0.$$

The first moment of $W(\eta_i,\eta_c)$ can be obtained according to the following relationships:

$$E[W(\eta_i,\eta_c)]$$

$$= -\frac{dw^*(s)}{ds}\Big|_{s=0} = \sum_{n=0}^{\infty}(1-\rho)\rho^n \frac{-dw_n^*(s)}{ds}\Big|_{s=0} = \sum_{n=0}^{\infty}(1-\rho)\rho^n E[Wn(\eta_i,\eta_c)]$$

$$= (1-\rho)\rho^0 E[W_0(\eta_i,\eta_c)] + (1-\rho)\sum_{n=1}^{\infty}\rho^n E[Wn(\eta_i,\eta_c)] = (1-\rho)\sum_{n=1}^{\infty}\rho^n E[Wn(\eta_i,\eta_c)]$$

$$= \frac{(1-\rho)}{\mu}\left[\sum_{n=1}^{\infty}\sum_{m=1}^{n}\sum_{k=m}^{\infty}\rho^{n+k-m}\eta_i^{\,k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})\right]$$

$$= \frac{(1-\rho)}{\mu}\left[\sum_{n=1}^{\infty}\sum_{m=1}^{n}\sum_{u=0}^{\infty}\rho^{n+u}\eta_i^{\,u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})\right]$$

$$= \frac{(1-\rho)}{\mu}\left[\sum_{n=1}^{\infty}\sum_{m=1}^{n}\rho^n\eta_i^{\,0}\prod_{j=0}^{-1}(1-(1-\eta_c)^{m+j}) + \sum_{n=1}^{\infty}\sum_{m=1}^{n}\sum_{u=1}^{\infty}\rho^{n+u}\eta_i^{\,u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})\right]$$

$$= \frac{(1-\rho)}{\mu}\left[\sum_{n=1}^{\infty}\sum_{m=1}^{n}\rho^n + \sum_{n=1}^{\infty}\sum_{m=1}^{n}\sum_{u=1}^{\infty}\rho^{n+u}\eta_i^{\,u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})\right]$$

$$= \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}n\rho^n + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{m=1}^{n}\sum_{u=1}^{\infty}\rho^{n+u}\eta_i^{\,u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})$$

$$= \frac{(1-\rho)}{\mu}\left(\frac{\rho}{(1-\rho)^2}\right) + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{m=1}^{n}\sum_{u=1}^{\infty}\rho^{n+u}\eta_i^{\,u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})$$

$$= \frac{\rho}{\mu-\lambda} + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{m=1}^{n}\sum_{u=1}^{\infty}\rho^{n+u}\eta_i^{\,u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})$$

$$= \frac{\rho}{\mu-\lambda} + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^{\,k}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right]. \quad \text{(B.8)}$$

This completes the proof of equation (4.2.2).

B.3. The proof of Corollary 4.2.1

Corollary 4.2.1 includes the results for the mean and variance of $W(\eta_i,\eta_c)$ at the boundary points. For the mean of $W(\eta_i,\eta_c)$, remember from equation (4.2.2) that

$$E[W(\eta_i,\eta_c)] = \frac{\rho}{\mu-\lambda} + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^{\,k}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right].$$

The mean of $W(\eta_i,\eta_c)$ at boundary points can be obtained from equation (4.2.2) according to the following relationships:

$$E[W(\eta_i,0)] = \frac{\rho}{\mu-\lambda} + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^{\,k}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-0)^{m+j})\right]$$

$$= \frac{\rho}{\mu-\lambda} + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^{\,k}[0] = \frac{\rho}{\mu-\lambda}. \qquad \text{(B.9)}$$

$$E[W(\eta_i,1)] = \frac{\rho}{\mu-\lambda} + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^{\,k}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-1)^{m+j})\right]$$

$$= \frac{\rho}{\mu-\lambda} + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}n\rho^{n}(\rho\eta_i)^k = \frac{\rho}{\mu-\lambda} + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}n\rho^{n}\sum_{k=1}^{\infty}(\rho\eta_i)^k$$

$$= \frac{\rho}{\mu-\lambda} + \frac{(1-\rho)}{\mu} * \frac{\rho\eta_i}{1-\rho\eta_i}\sum_{n=1}^{\infty}n\rho^{n} = \frac{\rho}{\mu-\lambda} + \frac{(1-\rho)}{\mu} * \frac{\rho^2\eta_i}{(1-\rho\eta_i)(1-\rho)^2}$$

$$= \frac{\rho}{\mu-\lambda} + \frac{(1-\rho)}{\mu} * \frac{\rho^2\eta_i}{(1-\rho\eta_i)(1-\rho)^2} = \frac{\rho}{\mu-\lambda}\left(\frac{1}{1-\rho\eta_i}\right). \qquad \text{(B.10)}$$

$$E[W(0,\eta_c)] = \frac{\rho}{\mu-\lambda} + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}(0)^k\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right]$$

$$= \frac{\rho}{\mu - \lambda}. \tag{B.11}$$

$$E[W(1,\eta_c)] = \frac{\rho}{\mu - \lambda} + \frac{(1-\rho)}{\mu} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \rho^{n+k} \left[\sum_{m=1}^{n} \prod_{j=0}^{k-1} (1-(1-\eta_c)^{m+j}) \right]. \tag{B.12}$$

The variance of $W(\eta_i, \eta_c)$ at boundary points can be obtained according to the following relationships:

$$Var[W(0,\eta_c)] = E\left(W^2(0,\eta_c)\right) - E^2\left(W(0,\eta_c)\right)$$

$$= \sum_{n=0}^{\infty} (1-\rho)\rho^n E[W_n^2(0,\eta_c)] - \left(\frac{\rho}{\mu - \lambda}\right)^2; \tag{B.13}$$

from equation (4.1.12), it is easy to find that

$$\begin{cases} E[W_n^2(0,\eta_c)] = \dfrac{2}{\mu} \displaystyle\sum_{m=1}^{n} \sum_{k=m}^{\infty} (0)^{k-m} \Big(\prod_{j=0}^{k-m-1} (1-(1-\eta_c)^{m+j}) \Big) E[W_k(0,\eta_c)]; \\ \\ E[W_k(0,\eta_c)] = \dfrac{1}{\mu} \displaystyle\sum_{m=1}^{k} \sum_{k=m}^{\infty} (0)^{k-m} \prod_{j=0}^{k-m-1} (1-(1-\eta_c)^{m+j}); \end{cases} \tag{B.14}$$

for $k = m$,

$$E[W_k(0,\eta_c)] = \frac{1}{\mu} \sum_{m=1}^{k} (0)^0 \prod_{j=0}^{-1} (1-(1-\eta_c)^{m+j}) = \frac{1}{\mu} \sum_{m=1}^{k} 1 = \frac{k}{\mu}; \tag{B.15}$$

By inserting equation (B.15) in equation (B.14) and setting $k = m$, the following result for $E[W_n^2(0,\eta_c)]$ can be obtained:

$$E[W_n^2(0,\eta_c)] = \frac{2}{\mu^2} \sum_{m=1}^{n} m(0)^0 \Big(\prod_{j=0}^{-1} (1-(1-\eta_c)^{m+j}) \Big)$$

$$= \frac{2}{\mu^2} \sum_{m=1}^{n} m = \frac{2n(n+1)}{2\mu^2} = \frac{n^2+n}{\mu^2}. \tag{B.16}$$

By inserting equation (B.16) in equation (B.13), the following result for $E\left(W^2(0,\eta_c)\right)$ can be obtained:

$$E\left(W^2(0,\eta_c)\right)=\sum_{n=0}^{\infty}(1-\rho)\rho^n\,\frac{n^2+n}{\mu^2}=\frac{1-\rho}{\mu^2}\left[\sum_{n=0}^{\infty}n\rho^n+\sum_{n=0}^{\infty}n^2\rho^n\right]$$

$$=\frac{1-\rho}{\mu^2}\left[\frac{\rho}{(1-\rho)^2}+\frac{\rho(1+\rho)}{(1-\rho)^3}\right]=\frac{2\rho}{\mu^2(1-\rho)^2}. \tag{B.17}$$

Finally, by inserting equation (B.17) in equation (B.13), it is easy to find $Var\,[W\,(0,\eta_c)]$ as follows:

$$Var[W(0,\eta_c)]=E\left(W^2(0,\eta_c)\right)-E^2\left(W(0,\eta_c)\right)$$

$$=\frac{2\rho}{\mu^2(1-\rho)^2}-\left(\frac{\rho}{\mu-\lambda}\right)^2=\frac{\rho(2-\rho)}{(\mu-\lambda)^2}. \tag{B.18}$$

$Var[W(\eta_i,1)]$ can be obtained according to the following relationships:

$$Var[W(\eta_i,1)]=E\left(W^2(\eta_i,1)\right)-E^2\left(W(\eta_i,1)\right). \tag{B.19}$$

$$E\left(W^2(\eta_i,1)\right)=\sum_{n=0}^{\infty}(1-\rho)\rho^n E[W_n^2(\eta_i,1)]. \tag{B.20}$$

From equation (4.1.12), $E[W_n^2(\eta_i,1)]$ can be obtained as follows:

$$E[W_n^2(\eta_i,1)]=\frac{2}{\mu}\sum_{m=1}^{n}\sum_{k=m}^{\infty}(\rho\eta_i)^{k-m}\Big(\prod_{j=0}^{k-m-1}1\Big)E[W_k(\eta_i,1)] \tag{B.21}$$

$$=\frac{2}{\mu}\sum_{m=1}^{n}\sum_{k=m}^{\infty}(\rho\eta_i)^{k-m}E[W_k(\eta_i,1)].$$

$$E[W_k(\eta_i,1)]=\frac{1}{\mu}\sum_{m=1}^{k}\sum_{k=m}^{\infty}(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}1=\frac{1}{\mu}\sum_{m=1}^{k}\frac{1}{1-\rho\eta_i}=\frac{k}{\mu(1-\rho\eta_i)}. \tag{B.22}$$

By inserting equation (B.22) in equation (B.21) and setting $(u = k - m)$, the following relationships can be obtained:

$$
\begin{aligned}
E[W_n^2(\eta_i,1)] &= \frac{2}{\mu^2(1-\rho\eta_i)}\sum_{m=1}^{n}\sum_{k=m}^{\infty}(\rho\eta_i)^{k-m}k = \frac{2}{\mu^2(1-\rho\eta_i)}\sum_{m=1}^{n}\sum_{u=0}^{\infty}(\rho\eta_i)^{u}(m+u) \\
&= \frac{2}{\mu^2(1-\rho\eta_i)}\left[\sum_{m=1}^{n}\sum_{u=0}^{\infty}m(\rho\eta_i)^u + \sum_{m=1}^{n}\sum_{u=0}^{\infty}u(\rho\eta_i)^u\right] \\
&= \frac{2}{\mu^2(1-\rho\eta_i)}\left[\sum_{m=1}^{n}\sum_{u=0}^{\infty}m(\rho\eta_i)^u + \sum_{m=1}^{n}\sum_{u=0}^{\infty}u(\rho\eta_i)^u\right] \\
&= \frac{2}{\mu^2(1-\rho\eta_i)}\left[\frac{n^2+n}{2(1-\rho\eta_i)} + \frac{n\rho\eta_i}{(1-\rho\eta_i)^2}\right] = \frac{(n^2+n)(1-\rho\eta_i)+2n\rho\eta_i}{\mu^2(1-\rho\eta_i)^3}.
\end{aligned}
\tag{B.23}
$$

By inserting equation (B.23) in equation (B.20), $E\left(W^2(\eta_i,1)\right)$ can be obtained according to the following relationships:

$$
\begin{aligned}
&E\left(W^2(\eta_i,1)\right) \\
&= \sum_{n=0}^{\infty}(1-\rho)\rho^n\left[\frac{(n^2+n)(1-\rho\eta_i)+2n\rho\eta_i}{\mu^2(1-\rho\eta_i)^3}\right] = \sum_{n=0}^{\infty}(1-\rho)\rho^n\left[\frac{n^2(1-\rho\eta_i)+n(1+\rho\eta_i)}{\mu^2(1-\rho\eta_i)^3}\right] \\
&= \frac{1}{\mu^2(1-\rho\eta_i)^3}\left[\sum_{n=0}^{\infty}(1-\rho)\rho^n n^2(1-\rho\eta_i) + \sum_{n=0}^{\infty}(1-\rho)\rho^n n(1+\rho\eta_i)\right] \\
&= \frac{1}{\mu^2(1-\rho\eta_i)^3}\left[(1-\rho)(1-\rho\eta_i)\sum_{n=0}^{\infty}\rho^n n^2 + (1-\rho)(1+\rho\eta_i)\sum_{n=0}^{\infty}\rho^n n\right] \\
&= \frac{1}{\mu^2(1-\rho\eta_i)^3}\left[(1-\rho)(1-\rho\eta_i)\frac{\rho(1+\rho)}{(1-\rho)^3} + (1-\rho)(1+\rho\eta_i)\frac{\rho}{(1-\rho)^2}\right] \\
&= \frac{1}{\mu^2(1-\rho\eta_i)^3}\left[\frac{\rho(1+\rho)(1-\rho\eta_i)+\rho(1-\rho)(1+\rho\eta_i)}{(1-\rho)^2}\right] \\
&= \frac{1}{\mu^2(1-\rho\eta_i)^3}\left[\frac{2\rho(1-\rho^2\eta_i)}{(1-\rho)^2}\right].
\end{aligned}
\tag{B.24}
$$

Finally, by inserting equation (B.24) in equation (B.19), it is easy to find $Var[W(\eta_i,1)]$ as follows:

$$Var[W(\eta_i,1)]$$

$$=\frac{1}{\mu^2(1-\rho\eta_i)^3}\left[\frac{2\rho(1-\rho^2\eta_i)}{(1-\rho)^2}\right]-\left((\frac{\rho}{\mu-\lambda})(\frac{1}{1-\rho\eta_i})\right)^2$$

$$=\frac{2\rho(1-\rho^2\eta_i)}{\mu^2(1-\rho)^2(1-\rho\eta_i)^3}-\frac{\rho^2}{(\mu-\lambda)^2(1-\rho\eta_i)^2}$$

$$=\frac{2\rho(1-\rho^2\eta_i)-\rho^2(1-\rho\eta_i)}{(\mu-\lambda)^2(1-\rho\eta_i)^3}=\frac{\rho(2-\rho-\rho^2\eta_i)}{(\mu-\lambda)^2(1-\rho\eta_i)^3}. \tag{B.25}$$

This completes the proof of Corollary 4.2.1.

B.4. The proof of Corollary 4.2.2

Remember from equation (4.2.4) that

$$\frac{\partial w(\eta_i,\eta_c)}{\partial\eta_i}=\frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}k\rho^{n+k}\eta_i^{k-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right].$$

Then, at $\eta_i=0$, it is easy to find that

$$\frac{\partial w(\eta_i,\eta_c)}{\partial\eta_i}\Big|_{\eta_i=0}=\frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}k\rho^{n+k}(0)^{k-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right]. \tag{B.26}$$

Equation (B.26) is equal to zero unless $k=1$. Then, the following relationships can be obtained:

$$\frac{\partial w(\eta_i,\eta_c)}{\partial\eta_i}\Big|_{\eta_i=0}=\frac{(1-\rho)\rho}{\mu}\sum_{n=1}^{\infty}\rho^n\sum_{m=1}^{n}(1-(1-\eta_c)^m)$$

$$=\frac{(1-\rho)\rho}{\mu}\sum_{m=1}^{\infty}\rho^n\sum_{n=m}^{\infty}(1-(1-\eta_c)^m)$$

$$= \frac{(1-\rho)\rho}{\mu} \sum_{m=1}^{\infty} (1-(1-\eta_c)^m) \sum_{n=m}^{\infty} \rho^n$$

$$= \frac{(1-\rho)\rho}{\mu} \sum_{m=1}^{\infty} (1-(1-\eta_c)^m) \frac{\rho^m}{(1-\rho)}$$

$$= \frac{\rho}{\mu} \sum_{m=1}^{\infty} \rho^m (1-(1-\eta_c)^m) = \frac{\rho}{\mu} \left[\sum_{m=1}^{\infty} \rho^m - \sum_{m=1}^{\infty} \left(\rho(1-\eta_c)\right)^m \right]$$

$$= \frac{\rho}{\mu} \left[\frac{\rho}{1-\rho} - \frac{\rho(1-\eta_c)}{1-\rho(1-\eta_c)} \right] = \frac{\rho^2(1-\rho(1-\eta_c)) - \rho^2(1-\rho)(1-\eta_c)}{\mu(1-\rho)(1-\rho(1-\eta_c))}$$

$$= \frac{\rho^2 - \rho^3 + \rho^3\eta_c - \rho^2 + \rho^3 + \rho^2\eta_c - \rho^3\eta_c}{\mu(1-\rho)(1-\rho(1-\eta_c))}$$

$$= \frac{\rho^2 \eta_c}{\mu(1-\rho)(1-\rho(1-\eta_c))}. \qquad \text{(B.27)}$$

$$\frac{\partial w(\eta_i, \eta_c)}{\partial \eta_i}\Big|_{\eta_i=1} = \frac{(1-\rho)}{\mu} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} k \rho^{n+k} (1)^{k-1} \sum_{m=1}^{n} \prod_{j=0}^{k-1} (1-(1-\eta_c)^{m+j})$$

$$= \frac{(1-\rho)}{\mu} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} k \rho^{n+k} \sum_{m=1}^{n} \prod_{j=0}^{k-1} (1-(1-\eta_c)^{m+j}). \qquad \text{(B.28)}$$

Remember from equation (4.2.4) that

$$\frac{\partial w(\eta_i, \eta_c)}{\partial \eta_c} = \frac{(1-\rho)}{\mu} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \rho^{n+k} \eta_i^k \left[\sum_{m=1}^{n} \sum_{i=0}^{k-1} (m+i)(1-\eta_c)^{m+i-1} \prod_{j=0; j \neq i}^{k-1} (1-(1-\eta_c)^{m+j}) \right].$$

Then, at $\eta_c = 0$, it is easy to find that

$$\frac{\partial w(\eta_i, \eta_c)}{\partial \eta_c}\Big|_{\eta_c=0} = \frac{(1-\rho)}{\mu} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \rho^{n+k} \eta_i^k \left[\sum_{m=1}^{n} \sum_{i=0}^{k-1} (m+i)(1-0)^{m+i-1} \prod_{j=0; j \neq i}^{k-1} (1-(1-0)^{m+j}) \right]$$

$$= \frac{(1-\rho)}{\mu} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \rho^{n+k} \eta_i^k \left[\sum_{m=1}^{n} \sum_{i=0}^{k-1} (m+i) \prod_{j=0; j \neq i}^{k-1} (0) \right]. \tag{B.29}$$

Equation (B.29) is equal to zero unless $k=1$, which leads to $i=0$ and $\prod_{j=1}^{0}(0)=1$. Then, the following relationships can be obtained:

$$\frac{\partial w(\eta_i, \eta_c)}{\partial \eta_c}\Big|_{\eta_c=0} = \frac{\rho \eta_i (1-\rho)}{\mu} \sum_{n=1}^{\infty} \rho^n \sum_{m=1}^{n} m = \frac{\rho \eta_i (1-\rho)}{2\mu} \sum_{n=1}^{\infty} \rho^n (n^2 + n)$$

$$= \frac{\rho \eta_i (1-\rho)}{2\mu} \left[\sum_{n=1}^{\infty} n^2 \rho^n + \sum_{n=1}^{\infty} n \rho^n \right] = \frac{\rho \eta_i (1-\rho)}{2\mu} \left[\frac{\rho(1+\rho)}{(1-\rho)^3} + \frac{\rho}{(1-\rho)^2} \right]$$

$$= \frac{\rho \eta_i}{2\mu} \left[\frac{\rho(1+\rho) + \rho(1-\rho)}{(1-\rho)^2} \right] = \frac{\rho \eta_i}{2\mu} \left[\frac{2\rho}{(1-\rho)^2} \right] = \frac{\rho^2 \eta_i}{\mu(1-\rho)^2}. \tag{B.30}$$

$$\frac{\partial w(\eta_i, \eta_c)}{\partial \eta_c}\Big|_{\eta_c=1} = \frac{(1-\rho)}{\mu} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \rho^{n+k} \eta_i^k \left[\sum_{m=1}^{n} \sum_{i=0}^{k-1} (m+i)(1-1)^{m+i-1} \prod_{j=0; j \neq i}^{k-1} (1-(1-1)^{m+j}) \right]$$

$$= \frac{(1-\rho)}{\mu} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \rho^{n+k} \eta_i^k \left[\sum_{m=1}^{n} \sum_{i=0}^{k-1} (m+i)(0)^{m+i-1} \prod_{j=0; j \neq i}^{k-1} 1 \right]. \tag{B.31}$$

Equation (B.31) is equal to zero unless $m=1$ and $i=0$. By considering this point in equation (B.31), it is easy to find that

$$\frac{\partial w(\eta_i, \eta_c)}{\partial \eta_c}\Big|_{\eta_c=1} = \frac{(1-\rho)}{\mu} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \rho^n (\rho \eta_i)^k [1] = \frac{(1-\rho)}{\mu} \sum_{n=1}^{\infty} \rho^n \sum_{k=1}^{\infty} (\rho \eta_i)^k$$

$$= \frac{(1-\rho)}{\mu} \sum_{n=1}^{\infty} \rho^n \frac{\rho \eta_i}{1-\rho \eta_i} = \frac{(1-\rho)\rho \eta_i}{\mu(1-\rho \eta_i)} \sum_{n=1}^{\infty} \rho^n = \frac{(1-\rho)\rho^2 \eta_i}{\mu(1-\rho \eta_i)(1-\rho)}$$

$$= \frac{\rho^2 \eta_i}{\mu(1-\rho \eta_i)}. \tag{B.32}$$

This completes the proof of Corollary 4.2.2.

B.5. The proof of Corollary 4.3.1

For $\rho < 1$, remember from equation (4.3.2) that

$$v^*(s) = 1-\rho+(1-\rho(1-\eta_c))\sum_{n=1}^{\infty}\rho^n(1-\eta_c)^{n-1}w_n^*(s).$$

The mean of $V(\eta_i,\eta_c)$ can be obtained according to the following relationships:

$$E[V(\eta_i,\eta_c)] = -\frac{dv^*(s)}{ds}\Big|_{s=0} = (1-\rho(1-\eta_c))\sum_{n=1}^{\infty}\rho^n(1-\eta_c)^{n-1}\frac{-dw_n^*(s)}{d_s}\Big|_{s=0}$$

$$= (1-\rho(1-\eta_c))\sum_{n=1}^{\infty}\rho^n(1-\eta_c)^{n-1}E[Wn(\eta_i,\eta_c)]$$

$$= \frac{(1-\rho(1-\eta_c))}{\mu}\sum_{n=1}^{\infty}\rho^n(1-\eta_c)^{n-1}\sum_{m=1}^{n}\sum_{k=m}^{\infty}(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})$$

$$= \frac{(1-\rho(1-\eta_c))}{\mu}\sum_{n=1}^{\infty}\rho^n(1-\eta_c)^{n-1}\sum_{m=1}^{n}\sum_{k=m}^{\infty}(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})$$

$$= \frac{(1-\rho(1-\eta_c))}{\mu}\sum_{n=1}^{\infty}\rho^n(1-\eta_c)^{n-1}\sum_{m=1}^{n}\sum_{u=0}^{\infty}(\rho\eta_i)^{u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})$$

$$= \frac{(1-\rho(1-\eta_c))}{\mu}\sum_{n=1}^{\infty}\rho^n(1-\eta_c)^{n-1}\sum_{m=1}^{n}\left(1+\sum_{u=1}^{\infty}(\rho\eta_i)^{u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})\right)$$

$$= \frac{(1-\rho(1-\eta_c))}{\mu}\sum_{n=1}^{\infty}\rho^n(1-\eta_c)^{n-1}\sum_{m=1}^{n}\left(1+\sum_{u=1}^{\infty}(\rho\eta_i)^{u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})\right)$$

$$= \begin{bmatrix} \dfrac{\rho(1-\rho(1-\eta_c))}{\mu}\displaystyle\sum_{n=1}^{\infty}n\big(\rho(1-\eta_c)\big)^{n-1} \\ +\dfrac{(1-\rho(1-\eta_c))}{\mu}\displaystyle\sum_{n=1}^{\infty}\rho^n(1-\eta_c)^{n-1}\sum_{m=1}^{n}\left(\sum_{u=1}^{\infty}(\rho\eta_i)^{u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})\right) \end{bmatrix}$$

$$=\left[\begin{array}{l}\dfrac{(1-\rho(1-\eta_c))}{\mu}*\dfrac{\rho}{\left(1-\rho(1-\eta_c)\right)^2} \\ +\dfrac{(1-\rho(1-\eta_c))}{\mu}\sum_{n=1}^{\infty}\rho^n(1-\eta_c)^{n-1}\sum_{u=1}^{\infty}(\rho\eta_i)^u\left(\sum_{m=1}^{n}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})\right)\end{array}\right]$$

$$=\frac{\rho}{\mu-\lambda(1-\eta_c)}+\frac{(1-\rho(1-\eta_c))}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(1-\eta_c)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right].$$

For $\rho \geq 1$, remember from equation (4.3.3) that

$$v^*(s)=\sum_{n=1}^{\infty}(1-\eta_c)^{n-1}\eta_c w_n^*(s), \qquad \rho \geq 1, \eta_i\rho < 1.$$

The mean of $V(\eta_i,\eta_c)$ can be obtained according to the following relationships:

$$E[V(\eta_i,\eta_c)]=-\frac{dv^*(s)}{ds}\Big|_{s=0}=\sum_{n=1}^{\infty}(1-\eta_c)^{n-1}\eta_c\left(\frac{-dw_n^*(s)}{d_s}\Big|_{s=0}\right)$$

$$=\sum_{n=1}^{\infty}(1-\eta_c)^{n-1}\eta_c E[Wn(\eta_i,\eta_c)]$$

$$=\frac{1}{\mu}\sum_{n=1}^{\infty}(1-\eta_c)^{n-1}\eta_c\sum_{m=1}^{n}\sum_{k=m}^{\infty}(\rho\eta_i)^{k-m}\prod_{j=0}^{k-m-1}(1-(1-\eta_c)^{m+j})$$

$$=\frac{1}{\mu}\sum_{n=1}^{\infty}(1-\eta_c)^{n-1}\eta_c\sum_{m=1}^{n}\sum_{u=0}^{\infty}(\rho\eta_i)^{u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j}), \tag{B.32}$$

By setting $u=k-m$ in equation (B.32), the following relationships are obtained:

$$=\frac{1}{\mu}\sum_{n=1}^{\infty}(1-\eta_c)^{n-1}\eta_c\sum_{m=1}^{n}\left(1+\sum_{u=1}^{\infty}(\rho\eta_i)^{u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})\right)$$

$$=\frac{1}{\mu}\sum_{n=1}^{\infty}(1-\eta_c)^{n-1}\eta_c\sum_{m=1}^{n}1+\frac{1}{\mu}\sum_{n=1}^{\infty}(1-\eta_c)^{n-1}\eta_c\sum_{m=1}^{n}\sum_{u=1}^{\infty}(\rho\eta_i)^{u}\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})$$

$$=\frac{\eta_c}{\mu}\sum_{n=1}^{\infty}n(1-\eta_c)^{n-1}+\frac{\eta_c}{\mu}\sum_{n=1}^{\infty}(1-\eta_c)^{n-1}\sum_{m=1}^{n}\sum_{u=1}^{\infty}(\rho\eta_i)^u\prod_{j=0}^{u-1}(1-(1-\eta_c)^{m+j})$$

$$=\frac{1}{\eta_c\mu}+\frac{\eta_c}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(\rho\eta_i)^k(1-\eta_c)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right]. \quad \text{(B.33)}$$

Expressions in equations (4.3.7), (4.3.8), (4.3.9) and (4.3.10) can be obtained from equation (4.3.5) according to the following calculations:

for $\rho<1$, remember from equation (4.3.5) that

$$E[V(\eta_i,\eta_c)]$$
$$=\frac{\rho}{\mu-\lambda(1-\eta_c)}+\frac{(1-\rho(1-\eta_c))}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(1-\eta_c)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right].$$

For the mean of the waiting time at boundary points, the following relationships can be obtained using equation (4.3.5):

$$E[V(\eta_i,0)]=\frac{\rho}{\mu-\lambda}+\frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(1-0)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-0)^{m+j})\right]$$

$$=\frac{\rho}{\mu-\lambda}+\frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(1-0)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(0)\right]=\frac{\rho}{\mu-\lambda}. \quad \text{(B.34)}$$

$$E[V(\eta_i,1)]=\frac{\rho}{\mu-\lambda(1-1)}+\frac{(1-\rho(1-1))}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(1-1)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-1)^{m+j})\right]$$

$$=\frac{\rho}{\mu}+\frac{1}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(0)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1)\right]=\frac{\rho}{\mu}+\frac{1}{\mu}\sum_{k=1}^{\infty}\rho^{k+1}\eta_i^k(0)^0\left[\sum_{m=1}^{1}\prod_{j=0}^{k-1}(1)\right]$$

$$=\frac{\rho}{\mu}+\frac{\rho}{\mu}\sum_{k=1}^{\infty}(\rho\eta_i)^k=\frac{\rho}{\mu}+\frac{\rho}{\mu}*\frac{\rho\eta_i}{1-\rho\eta_i}=\frac{\rho}{\mu}\left(1+\frac{\rho\eta_i}{1-\rho\eta_i}\right)=\frac{\rho}{\mu-\lambda\eta_i}. \quad \text{(B.35)}$$

For $\rho \geq 1$, remember from equation (4.3.5) that

$$E[V(\eta_i,\eta_c)] = \frac{1}{\eta_c\mu} + \frac{\eta_c}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(\rho\eta_i)^k(1-\eta_c)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right].$$

For the mean of the waiting time at the boundary points, the following relationships can be obtained using equation (4.3.5):

$$E[V(\eta_i,0)] = \frac{1}{\eta_c\mu} + \frac{0}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(\rho\eta_i)^k(1-0)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-0)^{m+j})\right] = +\infty. \quad \text{(B.36)}$$

$$E[V(\eta_i,1)] = \frac{1}{\mu} + \frac{1}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(\rho\eta_i)^k(1-1)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-1)^{m+j})\right]$$

$$= \frac{1}{\mu} + \frac{1}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(\rho\eta_i)^k(0)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}1\right] = \frac{1}{\mu} + \frac{1}{\mu}\sum_{k=1}^{\infty}(\rho\eta_i)^k(0)^0\left[\sum_{m=1}^{1}\prod_{j=0}^{k-1}1\right]$$

$$= \frac{1}{\mu} + \frac{1}{\mu}\sum_{k=1}^{\infty}(\rho\eta_i)^k = \frac{1}{\mu} + \frac{\rho\eta_i}{\mu(1-\rho\eta_i)} = \frac{1}{\mu}(1+\frac{\rho\eta_i}{1-\rho\eta_i}) = \frac{1}{\mu-\lambda\eta_i}. \quad \text{(B.37)}$$

$$E[V(0,\eta_c)] = \frac{1}{\eta_c\mu} + \frac{\eta_c}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(0)^k(1-\eta_c)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right] = \frac{1}{\eta_c\mu}. \quad \text{(B.38)}$$

B.6. The proof of equations (4.3.11)

For $\rho<1$, remember from equation (4.3.5) that

$$E[V(\eta_i,\eta_c)]$$
$$= \frac{\rho}{\mu-\lambda(1-\eta_c)} + \frac{(1-\rho(1-\eta_c))}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(1-\eta_c)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right].$$

$$\frac{\partial E[V(\eta_i,\eta_c)]}{\partial \eta_c}$$

$$= \begin{cases} \dfrac{-\lambda\rho}{[\mu-\lambda(1-\eta_c)]^2} + \dfrac{\rho}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(1-\eta_c)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right] \\ -\dfrac{(1-\rho(1-\eta_c))}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(n-1)(1-\eta_c)^{n-2}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right] \\ +\dfrac{(1-\rho(1-\eta_c))}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(1-\eta_c)^{n-1}\left[\sum_{m=1}^{n}\sum_{i=0}^{k-1}(m+i)(1-\eta_c)^{m+i-1}\prod_{\substack{j=0\\ j\neq i}}^{k-1}(1-(1-\eta_c)^{m+j})\right], \end{cases} \tag{B.39}$$

$$= \begin{cases} \dfrac{-\lambda\rho}{[\mu-\lambda]^2} + \dfrac{\rho}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(0)\right] \\ -\dfrac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(n-1)\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(0)\right] \\ +\dfrac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k\left[\sum_{m=1}^{n}\sum_{i=0}^{k-1}(m+i)(1-0)^{m+i-1}\prod_{\substack{j=0\\ j\neq i}}^{k-1}(0)\right], \end{cases} \tag{B.40}$$

$$= \frac{-\lambda\rho}{[\mu-\lambda]^2} + 0 - 0 + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k\left[\sum_{m=1}^{n}\sum_{i=0}^{k-1}(m+i)\prod_{\substack{j=0\\ j\neq i}}^{k-1}(0)\right]. \tag{B.41}$$

Equation (B.41) is equal to zero unless $k=1$. When $k=1$, then $i=0$ and $\prod_{j=1}^{0}(0)=1$ by convention. By considering this point in equation (B.41), it is easy to find that

$$\frac{\partial E[V(\eta_i,\eta_c)]}{\partial \eta_c}\Big|_{\eta_c=0} = \frac{-\lambda\rho}{[\mu-\lambda]^2} + \frac{(1-\rho)}{\mu}\sum_{n=1}^{\infty}\rho^{n+1}\eta_i\left[\sum_{m=1}^{n} m\prod_{j=1}^{0}(0)\right]$$

$$= \frac{-\lambda\rho}{[\mu-\lambda]^2} + \frac{\rho\eta_i(1-\rho)}{\mu}\sum_{n=1}^{\infty}\rho^{n}\left[\frac{n^2+n}{2}\right]$$

$$= \frac{-\lambda\rho}{[\mu-\lambda]^2} + \frac{\rho\eta_i(1-\rho)}{2\mu}\left[\sum_{n=1}^{\infty} n^2\rho^n + \sum_{n=1}^{\infty} n\rho^n\right]$$

$$= \frac{-\lambda\rho}{[\mu-\lambda]^2} + \frac{\rho\eta_i(1-\rho)}{2\mu}\left[\frac{\rho(1+\rho)}{(1-\rho)^3} + \frac{\rho}{(1-\rho)^2}\right]$$

$$= \frac{-\lambda\rho}{[\mu-\lambda]^2} + \frac{\rho\eta_i}{2\mu}\left[\frac{2\rho}{(1-\rho)^2}\right] = \frac{-\lambda\rho}{\mu^2(1-\rho)^2} + \frac{\rho^2\eta_i}{\mu(1-\rho)^2} = -\frac{\rho^2(1-\eta_i)}{\mu(1-\rho)^2}.$$

$$\frac{\partial E[V(\eta_i,\eta_c)]}{\partial \eta_c}\Big|_{\eta_c=1}$$

$$= \begin{cases} \dfrac{-\lambda\rho}{[\mu-\lambda(1-1)]^2} + \dfrac{\rho}{\mu}\displaystyle\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(1-1)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-1)^{m+j})\right] \\ -\dfrac{(1-\rho(1-1))}{\mu}\displaystyle\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(n-1)(1-1)^{n-2}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-1)^{m+j})\right] \\ +\dfrac{(1-\rho(1-1))}{\mu}\displaystyle\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(1-1)^{n-1}\left[\sum_{m=1}^{n}\sum_{i=0}^{k-1}(m+i)(1-1)^{m+i-1}\prod_{\substack{j=0\\ j\neq i}}^{k-1}(1-(1-1)^{m+j})\right] \end{cases} \tag{B.42}$$

$$= \begin{cases} \dfrac{-\lambda\rho}{\mu^2} + \dfrac{\rho}{\mu}\displaystyle\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(0)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}1\right] \\ -\dfrac{1}{\mu}\displaystyle\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(n-1)(0)^{n-2}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}1\right] \\ +\dfrac{1}{\mu}\displaystyle\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\rho^{n+k}\eta_i^k(0)^{n-1}\left[\sum_{m=1}^{n}\sum_{i=0}^{k-1}(m+i)(0)^{m+i-1}\prod_{\substack{j=0\\ j\neq i}}^{k-1}1\right] \end{cases} \tag{B.43}$$

$$= \frac{-\lambda\rho}{\mu^2} + \frac{\rho}{\mu}\sum_{k=1}^{\infty}\rho^{1+k}\eta_i^k - \frac{1}{\mu}\sum_{k=1}^{\infty}\rho^{2+k}\eta_i^k\left[\sum_{m=1}^{2}1\right] + \frac{1}{\mu}\sum_{k=1}^{\infty}\rho^{1+k}\eta_i^k$$

$$= \frac{-\rho^2}{\mu} + \frac{\rho^2}{\mu}\sum_{k=1}^{\infty}(\rho\eta_i)^k - \frac{2\rho^2}{\mu}\sum_{k=1}^{\infty}(\rho\eta_i)^k + \frac{\rho}{\mu}\sum_{k=1}^{\infty}(\rho\eta_i)^k$$

$$= \frac{-\rho^2}{\mu} + \sum_{k=1}^{\infty}(\rho\eta_i)^k\left(\frac{\rho^2}{\mu} - \frac{2\rho^2}{\mu} + \frac{\rho}{\mu}\right) = \frac{-\rho^2}{\mu} + \frac{\rho\eta_i}{1-\rho\eta_i}\left(\frac{-\rho^2+\rho}{\mu}\right)$$

$$= \frac{-\rho^2(1-\rho\eta_i) + \rho^2\eta_i - \rho^3\eta_i}{\mu(1-\rho\eta_i)} = \frac{-\rho^2(1-\eta_i)}{\mu(1-\rho\eta_i)}. \qquad \text{(B.44)}$$

This completes the proof of equation (4.3.11).

B.7. The proof of equations (4.3.12)

For $\rho \geq 1$, remember from equation (4.3.5) that

$$E[V(\eta_i,\eta_c)] = \frac{1}{\eta_c\mu} + \frac{\eta_c}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(\rho\eta_i)^k(1-\eta_c)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right].$$

$$\frac{\partial E[V(\eta_i,\eta_c)]}{\partial\eta_c}$$

$$= \begin{cases} -\frac{1}{\mu\eta_c^2} + \frac{1}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(\rho\eta_i)^k(1-\eta_c)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right] \\ -\frac{\eta_c}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(\rho\eta_i)^k(n-1)(1-\eta_c)^{n-2}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}(1-(1-\eta_c)^{m+j})\right] \\ +\frac{\eta_c}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(\rho\eta_i)^k(1-\eta_c)^{n-1}\left[\sum_{m=1}^{n}\sum_{i=0}^{k-1}(m+i)(1-\eta_c)^{m+i-1}\prod_{\substack{j=0\\ j\neq i}}^{k-1}(1-(1-\eta_c)^{m+j})\right]. \end{cases} \qquad \text{(B.45)}$$

$$\frac{\partial E[V(\eta_i,\eta_c)]}{\partial\eta_c}\Big|_{\eta_c=0} = -\infty. \qquad \text{(B.46)}$$

$$\frac{\partial E[V(\eta_i,\eta_c)]}{\partial \eta_c}\Big|_{\eta_c=1}$$

$$= \begin{cases} -\frac{1}{\mu}+\frac{1}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(\rho\eta_i)^k(0)^{n-1}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}1\right] \\ -\frac{1}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(\rho\eta_i)^k(n-1)(0)^{n-2}\left[\sum_{m=1}^{n}\prod_{j=0}^{k-1}1\right] \\ +\frac{1}{\mu}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(\rho\eta_i)^k(0)^{n-1}\left[\sum_{m=1}^{n}\sum_{i=0}^{k-1}(m+i)(0)^{m+i-1}\prod_{\substack{j=0\\ j\neq i}}^{k-1}1\right] \end{cases}$$

$$=-\frac{1}{\mu}+\frac{1}{\mu}\sum_{k=1}^{\infty}(\rho\eta_i)^k-\frac{2}{\mu}\sum_{k=1}^{\infty}(\rho\eta_i)^k+\frac{1}{\mu}\sum_{k=1}^{\infty}(\rho\eta_i)^k=-\frac{1}{\mu}. \qquad \text{(B.47)}$$

This completes the proof of equation (4.3.12).

B.8. The proof of equation (4.4.3)

$$Var(W_a(\eta_i,\eta_c)) = E\left(W_a^2(\eta_i,\eta_c)\right)-E^2\left(W_a(\eta_i,\eta_c)\right). \qquad \text{(B.48)}$$

The following relationship for $E\left(W_a^2(\eta_i,\eta_c)\right)$ can be obtained:

$$E\left(W_a^2(\eta_i,\eta_c)\right)=\int_0^{\infty}t^2dP\{W_a(\eta_i,\eta_c)<t\}$$

$$=\int_0^{\infty}t^2d\left\{\eta_iP\{V(\eta_i,\eta_c)<t\}+(1-\eta_i)P\{W(\eta_i,\eta_c)<t\}\right\}$$

$$=\eta_i\int_0^{\infty}t^2dP\{V(\eta_i,\eta_c)<t\}+(1-\eta_i)\int_0^{\infty}t^2dP\{W(\eta_i,\eta_c)<t\}$$

$$=\eta_iE\left(V^2(\eta_i,\eta_c)\right)+(1-\eta_i)E\left(W^2(\eta_i,\eta_c)\right); \qquad \text{(B.49)}$$

The following relationship for $E\left(W_a(\eta_i,\eta_c)\right)$ can be obtained:

$$E\left(W_a\,(\eta_i,\eta_c)\right)=\eta_i\int_0^{\infty} t\ \ dP\{V(\eta_i,\eta_c)<t\}+(1-\eta_i)\int_0^{\infty} t\ \ dP\{W(\eta_i,\eta_c)<t\}$$

$$=\eta_i E\left(V(\eta_i,\eta_c)\right)+(1-\eta_i)E\left(W(\eta_i,\eta_c)\right); \qquad \text{(B.50)}$$

By inserting equations (B.49) and (B.50) in equation (B.48), it is easy to find that

$$\begin{aligned}
&Var(W_a(\eta_i,\eta_c))\\
&=\eta_i E\left(V^2(\eta_i,\eta_c)\right)+(1-\eta_i)E\left(W^2(\eta_i,\eta_c)\right)-\left[\eta_i E\left(V(\eta_i,\eta_c)\right)+(1-\eta_i)E\left(W(\eta_i,\eta_c)\right)\right]^2\\
&=\eta_i\left[Var\left(V(\eta_i,\eta_c)\right)+E^2\left(V(\eta_i,\eta_c)\right)\right]+(1-\eta_i)\left[Var\left(W(\eta_i,\eta_c)\right)+E^2\left(W(\eta_i,\eta_c)\right)\right]\\
&\qquad\qquad-\left[\eta_i E\left(V(\eta_i,\eta_c)\right)+(1-\eta_i)E\left(W(\eta_i,\eta_c)\right)\right]^2\\
&=(1-\eta_i)Var(W(\eta_i,\eta_c))+\eta_i Var(V(\eta_i,\eta_c))\\
&\qquad\qquad+\eta_i(1-\eta_i)(E[W(\eta_i,\eta_c)]-E[V(\eta_i,\eta_c)])^2.
\end{aligned}$$

This completes the proof of equation (4.4.3).

Computer Programming

Waiting Time Analysis

```
#include <stdio.h>
#include <stdlib.h>
#include <math.h>

#define NofEta 1500
#define Nlimit 1500

FILE *cfPtr;

float lambda, mu, eta_i ,eta_c ,E_w[Nlimit] ,E_t[NofEta],E_t2[NofEta],E_w2[Nlimit],
Flag_1[Nlimit], Flag_2[Nlimit],
                    VarW, Ro, EW, EW2, EV, EV2, VarV, EWa1, EWa2, VarWa;
int N,i,j, Flag_Condition;
int k ;
```

// The following commands open the input file and scan the input parameters that are λ, μ, N, η_i and η_c, respectively.//

```
int main()

{
if ( ( cfPtr = fopen ("INPUT.txt", "r" ) ) == NULL)
                    printf( "File couldnot be opened\n");
          else {
                    fscanf(cfPtr, "%f%f%d", &lambda,&mu,&N);
                    fscanf(cfPtr, "%f%f", &eta_i,&eta_c);
```

// The following command computes the amount of E[T_N] according to equation (5.3).//

```
E_t[N]=(float)(1/(mu-lambda*eta_i));  // Equatin (5.3)//
```

// the following command computes the amount of E[T_n] for $n<N$, according to equation (5.4).//

```
for ( i=N-1; i>=1; i-- )
    E_t[i]=(float)( 1/mu + (lambda*eta_i*(1-pow((1-eta_c),i) ) /mu )*E_t[i+1] );
```

//The following command computes the amount of E[T^2_N] according to equation (5.3).//

```
E_t2[N]=(float)2*mu/pow((mu-lambda*eta_i),3);  // Equation (5.3)//
```

//The following command computes the amount of E[T^2_n] for $n<N$ according to equation (5.4).//

```
for (i=N-1; i>=1; i-- )
    E_t2[i]=(float)(2*E_t[i]+lambda*eta_i*(1- pow ( ( 1-
            eta_c),i))*(E_t2[i+1]+2*E_t[i]*E_t[i+1]))/mu;

if ( ( cfPtr = fopen( "OutPut.txt", "w" ) ) == NULL )

    printf( "File couldnot be opend\n" );

else {
```

// The following command computes the amount of $E(W_n(\eta_i,\eta_c))$ for $n<=N$ according to equation (5.5).//

```
E_w[0] = 0 ;
for ( i=1; i<=N; i++ )
    E_w[i] = (float) E_w[i-1] + E_t[i];
```

// The following command computes the amount of $E(W_n^2(\eta_i,\eta_c))$ for $n<=N$ according to equation (5.5).//

```
E_w2[0] = 0 ;
for ( i=1 ; i<=N; i++ )
    E_w2[i] = (float) E_w2[i-1]+E_t2[i]+(2*E_w[i-1]*E_t[i]) ;
```

// The following commands compute the amount of $E(W(\eta_i, \eta_c))$ according to equation (4.2.1).//

```
Ro = (float) lambda / mu;
EW= 0 ;
for ( i=1; i<=N; i++ )
    EW = (float) EW + (1-Ro)*pow ((Ro),i)*E_w[i] ;
```

// The following commands compute the amount of $E(W^2(\eta_i, \eta_c))$ according to equation (4.2.1).//

```
EW2 = 0 ;
for ( i=1; i<=N; i++ )
    EW2 = (float) EW2 + (1-Ro)*pow((Ro),i)*E_w2[i] ;
```

// The following commands compute the amount of $Var(W(\eta_i, \eta_c))$ .//

```
VarW = 0 ;
VarW = (float) EW2 - pow((EW),2) ;
```

// The following commands compute the amount of $E(V(\eta_i, \eta_c))$ and $E(V^2(\eta_i, \eta_c))$ according to equation (4.3.4).//

```
EV  =  0 ;
EV2 =  0 ;
if (Ro < 1 ){
    for ( i=1; i<=N; i++ )
        EV = (float)EV +  ( 1-Ro*(1-eta_c))*pow(Ro, i) *pow((1-
                                            eta_c),i-1)*E_w[i] ;
    for ( i=1; i<=N; i++ )
        EV2 = (float) EV2 +  ( 1-Ro*(1-eta_c))*pow(Ro, i) *pow((1-
            eta_c),i-1)*E_w2[i] ;
}
else{
```

```
                    for ( i=1; i<=N; i++ )
                         EV = (float) EV + pow(( 1-eta_c ),i-1)*eta_c*E_w[i] ;
                    for ( i=1; i<=N; i++ )
                         EV2 = (float) EV2 + pow(( 1-eta_c ),i-1)*eta_c*E_w2[i] ;
               }
```

// The following command computes the amount of $Var(V(\eta_i,\eta_c))$ .//

```
               VarV  = (float) EV2 - pow(EV,2) ;
```

// The following commands compute the amount of $E(W_a(\eta_i,\eta_c))$ according to equation (4.4.2) .//

```
               EWa2 = 0 ;
               for ( i=1; i<=N; i++ )
                    EWa2 =(float)  eta_i * EV + (1- eta_i) * EW ;
```

// The following commands compute the amount of $Var(W_a(\eta_i,\eta_c))$ according to equation (4.4.3) .//

```
               if (Ro < 1 ){
                    VarWa = (1-eta_i) *VarW+ eta_i* VarV+eta_i*(1-eta_i)*pow((EW –
                         EV),2) ;
                    fprintf(cfPtr,"VarWa = %f\n ",VarWa) ;
               }
               else
                    fprintf(cfPtr,"VarWa = ,Infinite\n ") ;
```

// The following commands print the results for the mean and variance of customers in the output file.//

```
               fprintf(cfPtr,"EW = %f\n ", EW) ;
               fprintf(cfPtr,"VarW = %f\n ", VarW) ;
               fprintf(cfPtr,"EV = %f\n ", EV) ;
               fprintf(cfPtr,"VarV = %f\n ",VarV) ;
               fprintf(cfPtr,"EWa2 = %f\n ",EWa2) ;
          }
          }
                         fclose (cfPtr);
                         return 0;
}
```